DOCUMENT RESUME

ED 285 516 HE 020 862

AUTHOR	Brown, George H.; Faupel, Elizabeth M.
TITLE	Postsecondary Assessment. Report of a Planning Conference (Washington, D.C., November 20, 1986).
INSTITUTION	Center for Education Statistics (OERI/ED), Washington, DC.
REPORT NO	CS-87-445
PUB DATE	87
NOTE	114p.; For reports from related conferences, see PS 016 941 and TM 870 644.
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402.
PUB TYPE	Collected Works - Conference Proceedings (021) -- Viewpoints (120)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Academic Achievement; Academic Persistence; Agency Role; Alumni; College Graduates; *Educational Assessment; Federal Government; *Government Role; *Government School Relationship; *Outcomes of Education; *Postsecondary Education; State Government; *Student Evaluation
IDENTIFIERS	*Center for Education Statistics

ABSTRACT

A report of a 1986 planning conference on postsecondary assessment is presented along with four commissioned papers from the conference and their summaries. Titles and authors of the papers are as follows: "Postsecondary Assessment: The State of the Art" (Leonard L. Baird); "Assessing Student Progress in College: A Process-Oriented Approach to Assessment of Student Learning in Postsecondary Settings" (Paul R. Pintrich); "Assessment of College Outcomes" (Aubrey Forrest); and "The State of the States in Postsecondary Assessment" (Edward A. Morante). An edited version of the general discussion that ended the conference is also provided. The objectives of the discussion were to summarize the conference and elicit recommendations for the Center for Education Statistics (CES). The final section of the report considers implications for the Center. A paper prepared by Dennis Jones, "Postsecondary Education Assessment: The Role of the Center for Education Statistics," synthesizes the conference. In addition, comment on the conference is provided by David A. Sweet, Director, Education Outcomes Division, CES. A list of conference participants and biographical sketches of speakers are appended. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Postsecondary Assessment
Report of a Planning Conference
November 20, 1986

Prepared by
George H. Brown
with
Elizabeth M. Faupel

U.S. Department of Education
William J. Bennett
Secretary

Office of Educational Research and Improvement
Chester E. Finn, Jr.
Assistant Secretary

Center for Education Statistics
Emerson J. Elliott
Director

Information Services
Edwin S. Darrell
Director

Center for Education Statistics

"The purpose of the Center shall be to collect and disseminate statistics and other data related to education in the United States and in other nations. The Center shall . . . collect, collate, and from time to time, report full and complete statistics on the conditions of education in the United States; conduct and publish reports on specialized analyses of the meaning and significance of such statistics; . . . and review and report on education activities in foreign countries,"—Section 406(b) of the General Education Provisions Act, as amended (20 U.S.C. 1221e-1).

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

CONTENTS

Chapter		Page
1.	Introduction	1
2.	Paper Summaries	5
3.	Summary of General Discussion	21
4.	Implications for the Center	24
	A. Postsecondary Assessment: the Role of the Center for Education Statistics, by Dennis P. Jones	27
	B. Comments, by David A. Sweet	43
Appendixes		
	A. Commissioned Papers	45
	(1) Postsecondary Assessment: The State of the Art, by Leonard L. Baird	47
	(2) Assessing Student Progress in College, by Paul R. Pintrich	67
	(3) Assessment of College Outcomes, by Aubrey Forrest	93
	(4) The State of the States in Postsecondary Assessment, by Edward A. Morante	101
	B. List of Conference Participants	113
	C. Biographical Sketches of Speakers	117

0 2,412 411

Chapter 1

Introduction

Assessment of educational attainment and the companion notion of accountability are widely discussed at many levels of education. If assessments can be made of the educational attainments of students from a particular educational system, then it may be possible to hold that system accountable for deficiencies in skills it produced in its graduates. In theory, at least, assessment data can serve a diagnostic function and identify the aspects of a system that need to be changed. Assessment might function as a "lever for change."

Educational assessment that is national in scope and intended to provide a sort of global indicator of U.S. education effectiveness, began with the National Assessment of Educational Progress (NAEP) in 1969. NAEP, which is sponsored by the U.S. Department of Education, periodically assessed the educational attainments of 9-, 13-, and 17-year olds, and occasionally, young adults. NAEP has not assessed educational attainment at the postsecondary level.

Should the quality of postsecondary education be assessed? Should the U.S. Department of Education play some role in postsecondary assessment? Questions of this sort have been raised in the world of higher education. On many occasions, U.S. Secretary of Education William J. Bennett has questioned whether colleges are fulfilling their obligations to students of ensuring that when they leave, it is as educated men and women. He directed Assistant Secretary Chester E. Finn, Jr. to explore the feasibility of various approaches to assessing the knowledge and skills possessed and acquired by college students. This conference on postsecondary assessment is an early step in responding to that directive.

Many State departments of education and institutions of higher education are developing or using some form of assessment system at the postsecondary level.

The term "postsecondary assessment" does not have a unique meaning. To some, it means the assessment of the skills and qualifications of entering freshmen--a sort of assessment of readiness for college, and indirectly, of the requirements for remedial instruction.

Sometimes it means the measurement of students' progress through college--of what they learned after 1, 2, or 3 years of college. The "rising junior" examination is an example of this concept of postsecondary assessment.

A third meaning, and perhaps the most widely shared, is a measurement of college outcomes--of what students learn by the time they graduate. Traditionally, colleges have evaluated

themselves and other institutions, by measures such as the faculty credentials, size of endowment, and the aptitude scores of entering freshmen. The third concept of postsecondary assessment stresses the results of education rather than what goes into it.

Selection of Panel Members

A distinguished panel was selected from a list of experts recommended to the Department by individuals who themselves had excellent credentials in postsecondary education and a strong interest in postsecondary assessment.

The panelists' charge was to address the following:

> What do students need to know and do to benefit from postsecondary education? How can student performance best be assessed at the postsecondary (entry, midterm, or completion) level? Is "value added" a useful ccncept for postsecondary outcomes and, if so, what implications could it have for the design of a longitudinal study? What sort of postsecondary assessment activities are underway in various departments of education at the State level? What can we learn from them that would be important to know in designing a longitudinal study?

Note at the planning stage of the conference, postsecondary assessment had a strong longitudinal studies orientation. By the time the conference took place, this orientation was no longer dominant. In the early stages of conference planning, the door was left open for considering postsecondary assessment in all its manifestations.

Four panelists were commissioned to present a paper on an area of postsecondary assessment. An expert was commissioned to serve as synthesizer and discussion leader. His task was to distill from the discussions a set of explicit recommendations as to what statistical activities, if any, the Center for Education Statistics should undertake in the general area of postsecondary assessment.

In discussions with each panel member, the objective of the conference was emphasized, to develop recommendations for the Center for Education Statistics's role in postsecondary assessment.

Plan of This Report

Chapter 2 gives summaries of the four commissioned papers. The four authors and their papers were

* The State of the Art in Postsecondary Assessment, by Leonard L. Baird, University of Kentucky.

* Assessing Student Progress in College: A Process-Oriented Approach to Assessment of Student Learning in Postsecondary Settings, by Paul R. Pintrich, University of Michigan.

* Assessment of College Outcomes, by Aubrey Forrest, American College Testing Program.

* The State of the States in Postsecondary Assessment, by Edward A. Morante, New Jersey Department of Higher Education.

Chapter 3 gives an edited version of the general discussion which was the final activity of the conference. Discussion leader was Dennis P. Jones, National Center for Higher Education Management Systems (NCHEMS). The objectives of the discussion were to summarize the conference and elicit recommendations for the Center.

Chapter 4, "Implications for the Center," comprises two parts. The first part is a paper prepared by Dennis Jones synthesizing the conference, "The Role of the Center for Education Statistics." The second part is comment on the conference by David A. Sweet, Director, Education Outcomes Division, Center for Education Statistics.

The appendix contains the four papers (the authors may have made slight revisions after presentation), a list of conference participants, and biographical sketches of panel members.

Chapter 2

Paper Summaries

This chapter presents summaries of the four papers commissioned for the conference. Each summary was prepared by the author.

Summary

Postsecondary Assessment: The State of the Art

Leonard L. Baird
University of Kentucky

Most discussions of "postsecondary assessment" focus on the measurement of students' knowledge and academic skills when they apply to or enter college, when they are in the midst of their studies, and at the end of their college careers. Although these assessments are important and involve many conceptual and technical problems, they are only part of the story. Postsecondary assessment, broadly conceived, includes information about a variety of aspects of postsecondary education that help us understand, monitor, and evaluate its processes and programs. It is important to realize that we cannot fully understand why students know what they know and have the skills they do without assessing and understanding these other aspects of postsecondary education.

To assist in seeing these other aspects, a "map" of postsecondary assessment is described. The "map" is not intended to be completely descriptive or imply causation. Rather, it is a device to show how various important parts of postsecondary education are integrated and flow into each other. It represents an attempt to put our various questions about postsecondary education into focus, and consider the availability of conceptual models, the identification of variables, the measurement of those variables, and the feasibility of obtaining those measures on a large scale basis. There are 20 points on the map which represent areas which require assessment:

(1) Precollege characteristics, including academic preparation, demographic and background characteristics;

(2) The transition from high school to college, including college choice and the influence of finances, access, and background on attendance;

(3) Adult entrance to postsecondary education;

(4) College characteristics, focusing on the differences among colleges and the significance of those differences;

(5) Nontraditional forms of postsecondary education, including their definition, scope, and quality;

(6) Within college experiences and their influence, particularly residential status and major;

(7) The influence of students' pre-collegiate characteristics upon their college experiences, for

example, the kinds of students who choose various majors;

(8) The influence of college characteristics upon college experiences, for example, whether student-faculty interaction is less frequent in large universities because of their organizational structures;

(9) The college environment, or the subjective nature of the college experience, including the sense of community and "involvement;"

(10) The influence of college structures and student characteristics on the environment;

(11) Accurate information on retention and attrition, a considerably more problematical and difficult area than is first apparent;

(12) Prediction and understanding of the retention/attrition process;

(13) College outcomes, focusing on those which reflect attainment of the goals of postsecondary education;

(14) College effects, or the influence of colleges and their programs on student outcomes;

(15) The transition to graduate or professional education;

(16) Assessment of the varieties and environments of graduate and professional schools;

(17) Attrition and retention in graduate and professional study;

(18) The competence and outcomes of students at the completion of graduate and professional school;

(19) Career or life success, in all their varieties and vagaries;

(20) The prediction or relation of life success to previous educational attainments and experiences.

Consideration of this map suggests several areas in which we need to improve our understanding and assessments. One concerns the increasing numbers of adults entering postsecondary education. Clearly we would benefit greatly by some large scale studies of their characteristics, motivations, learning styles and preparation. A second gap is graduate and professional education, where most of what we know is limited to the elite graduate and professional programs. There are many other gaps in our information.

Two relatively inexpensive ways to fill in some of these gaps are proposed. The first is an annual senior survey of national samples of college students, roughly analogous to the ACE/UCLA freshman surveys, which would allow us to chart trends in student career choices, plans, experiences in college, indebtedness, and evaluations of their colleges. The second is to make more use of the data sets of the Educational Testing Service and American College Testing Program, which routinely collect vast amounts of data on college applicants and college students.

Of more fundamental importance is the development of models to understand the processes of postsecondary education. An example is the area of student attrition and retention, which has been made much more coherent and understandable by the use of models developed by Spady and Tinto. Instead of an atheoretical shotgun approach, researchers using this model have been able to identify and assess the influence on attrition of student background, student goals and college experiences. They have also been able to better understand the nature of the college experience. Another example is an interpretive model of the changing meaning of postsecondary education from opportunities to ultimatums. That is, what once was an almost guaranteed route to social and economic mobility is now a requirement for having the chance for mobility, or even for retaining one's current status. As these examples suggest, the "state of the art" in postsecondary assessment depends much more on our understanding than on our techniques. Advances will come more from the quality of our ideas than the quality of our methods.

Summary

Assessing Student Progress in College: A Process-Oriented Approach to Student Learning in Postsecondary Settings

Paul R. Pintrich
The University of Michigan

The role of assessment in education has a long history in educational research and policy (Haney, 1984). As Linn (1986a) points out, discussions of testing and assessment issues are ubiquitous in American education. Moreover, both Haney (1984) and Linn (1986a) note that past and current calls for educational reform have used test results to buttress arguments about the need for improvement in the educational system and also suggest that assessment can be one tool to facilitate change. In general, the calls for reform have been aimed at improving our elementary and secondary schools, but recently there has been increased interest and emphasis on improving postsecondary institutions (e.g. Boyer, forthcoming; Mortimer et al. 1984). Paralleling the elementary and secondary reports, the reports on postsecondary education have suggested that assessment can play an important role in improving our efforts at the college level.

There are, of course, a variety of issues to consider in discussing assessment at the postsecondary level. Alexander and Stark (1986) have suggested that there are eight issues to consider in any assessment program, but this paper focuses on two of the most important. The two are 1) What are the purposes of the assessment? and 2) How will the assessment data be used to improve college teaching and student learning? This paper presents a relatively circumscribed view of the purpose and use of assessment in postsecondary settings. Specifically, this paper develops a process-oriented view of assessment that assumes that one of the most important purposes of assessment is to improve teaching and learning in postsecondary settings. Given this purpose, assessment programs need to be closely linked to teaching and learning problems at the college level. Moreover, one of the best uses of assessment data is to focus on specific recommendations to improve the teaching and learning process. Accordingly, assessment in this view will have to concentrate on student, faculty, and classroom level data and the processes of learning and teaching. Although this type of data may not readily serve State or Federal policy interests, it should serve the interests of faculty and students. As Scarr (1981) has stated, "testing should always be used in the interests of the children tested." (P. 1159.) In addition, assessment should serve the interests of college faculty.

Following this assumption that assessment programs should focus on how assessment data can be used to improve teaching and learning at the local level, the paper suggests that assessment should focus on four domains of student outcomes. The four domains are 1) student knowledge and conceptualization of subject

matter, 2) students' learning strategies, 3) students' critical thinking, and 4) students' motivation, attitudes, and values towards lifelong learning. The paper reviews a variety of theoretical models that have addressed these four domains and suggests that a general information processing model is best suited to guide assessments efforts.

In addition, the paper attempts to encourage the development of local models of teaching and learning that parallel recent work by evaluation researchers. For example, Corday (1986) suggests that evaluators build conceptul models that specify student entry characteristics, treatment variables (including treatment fidelity and implementation concerns), mediating variables, and outcome variables. The specification of these models for postsecondary education assessment programs will help us avoid simplistic input-output models of college impact that, at worst, may be misleading, and, at best, do not provide useful information that can be readily used to improve teaching and learning.

The paper concludes with the following five points:

1) As Haney (1984), Linn (1986), and Sternberg (1985) have all noted, new assessment or testing programs must be more closely tied to instructional concerns as well as current theory on cognitive models of student learning. Accordingly, if researchers and policy makers are serious about the goal of improving postsecondary teaching and learning and plan to use assessment programs as one tool to reach this goal, then the assessment program must be linked to a strong, theoretical framework of student learning, motivation, and instruction. Accordingly, assessment programs that are basically atheoretical in nature, that include a variety of items tapping a diversity of student outcomes, and are not linked to concerns of teaching and learning, will not be very useful in improving teaching and learning in higher education. Although assessment programs that are atheoretical and based on general "dustbowl empiricism" methods may have some administrative utility, they generally do not provide information that is readily usable for instructional improvement efforts.

2) Although I believe the general information processing models presented in this paper are the most relevant, there may be other models that are useful in designing assessment programs. However, it is important to note that these other models should be concerned not only with the psychology of student learning, but also with the psychology of instruction.

3) As Corday (1986) has pointed out, simple input-output models are not adequate for most program evaluation tasks. Assessment programs need to focus not only on student outcomes, but also on instructional processes and other mediating constructs.

4) Given the goal of improving teaching and learning, the process-oriented model presented in this paper implies that many of the policy parameters and decisions discussed by Alexander and Stark (1986) will be focused at the local institutional or faculty level. Accordingly, colleges need to develop the internal capabilties to implement the types of assessment programs suggested here. For example, institutional researchers and faculty need to be trained in the development and use of these programs.

5) Finally, given this local policy focus, State and Federal policy makers may have only an indirect role to play in the development of the suggested assessment programs. As Bennett (1986) has pointed out, "The federal government cannot and should not play the primary role in the assessment of higher education." (P. iii.) Accordingly, the Federal Government should foster assessment activities that help to improve higher education by focusing the discussion on how assessment programs can be used to improve teaching and learning.

Summary

Assessment of College Outcomes

Aubrey Forrest
American College Testing Program

This brief paper suggests a role for the Federal Government in the assessment of college outcomes that envisions groups of cooperating colleges and universities reporting results of their data collection efforts to the Federal Government as group data. Not recommended here is any additional reporting by individual institutions directly to the Federal Government. This will protect the confidentiality of the individual institutions while providing data to the Nation through the Federal Government on the general state of American higher education as well as experimental efforts to improve undergraduate education.

It is anticipated that the cooperating groups of institutions will likely be formed by various types (large, small, public, private, two-year, four-year, etc.). They may correspond to existing State systems and/or private consortia. Funding for the cooperative data collection and reporting efforts may come from Federal, State or private sources. The reporting of such data to the Federal Government would be followed by dissemination of the data by the Federal Government to interested parties across the Nation.

The college outcome assessment data which the Federal Government should assist in collecting and disseminating should be that which the colleges and universities will find to be most critical to program improvement and that the general public will find to be most convincing in making appropriate judgments about the quality of American undergraduate education. In this light, three types of outcome data would appear to be most useful: (1) achievement test score gains, (2) student retention/persistence rates and (3) alumni satisfaction indices. The Federal Government should do what it can to assist and encourage the collection of these types of data by cooperating groups of colleges.

Lastly, it is suggested that the Federal Government actively promote a national agenda of undergraduate program changes that colleges and universities should implement on an experimental basis, with appropriate data collection procedures, to determine what strategies work best to increase institutional performance on the three indicators cited above.

Summary

The State of the States in Postsecondary Assessment

Edward A. Morante
New Jersey Department of Higher Education

Introduction

A series of reports on higher education have been issued in recent years addressing some aspect of a perceived need for reform. Many individual colleges and more than half of the States have begun initiatives to improve the quality of collegiate education using some form of assessment.

For this paper, "postsecondary assessment" falls into three categories:

* Developmental education (from assessing students' readiness for college to evaluating remedial programs).

* College-level learning (testing of students' learning at various points in their college career; sometimes referred to as "rising junior" or "gateway" exams, but need not be either).

* Outcomes assessment (comprehensive assessment programs evaluating areas beyond student learning and incorporating measures beyond tests).

State by State

Postsecondary assessment differs greatly from State to State. Many States have no initiative in this area although individual colleges within their borders may be actively involved in assessment activities. At the other end of the spectrum, several States have detailed mandates requiring colleges to participate participate in assessment.

Postsecondary assessment, especially at the State level, is a relatively new phenomenon with a recent flurry of activity. Any specific national overview of postsecondary assessment will probably be out of date, therefore, by the time it is distributed. With this caveat, the following provides a sampling of State-level activities in postsecondary assessment.

Connecticut. The State is requiring assessment but allows each institution to develop its own program, methods, and instruments with annual reporting to begin not before 1989-1990.

Florida. The CLAST (College Level Academic Skills Test) has been administered in Florida since 1982. Students in public colleges cannot receive an asociate degree or continue to the

junior year without passing this test of reading, writing, and computation. More recently, Florida has begun requiring basic skills testing at entry. Using one of four possible tests, students who do not meet standards must enroll in remedial courses.

Georgia. This State also has two statewide testing programs: at entry (for those below a cut-off score on the College Board's admission test) and at the end of the sophomore year (Regents tests) that must be passed to qualify for graduation.

Rhode Island. All three public institutions must develop comprehensive assessment programs under guidelines set by the Department of Higher Education office. Cyclical reports are expected with emphasis on the utilization of existing data and assessment mechanisms (i.e., accreditation self-study reports).

Tennessee. Performance-based funding has been used for educational involvement since 1979. Colleges meeting statewide guidelines receive funding beyond their normal allocation. The ACT COMP test, required of all students prior to graduation, is also used for institutional evaluations. More recently, the Board of Regents implemented a statewide system of basic skills testing at entry and remedial program evaluation.

Texas. A statewide faculty/administrative committee recommended basic skills testing of all students before passing them on to their junior year. No action has yet been taken on this recommendation.

New Jersey. This State has two separate programs, one of long standing and one under development:

* Basic Skills Assessment Program. Since 1978 all entering students in public colleges are required to take a State-developed basic skills test (New Jersey College Basic Skills Placement Test) before enrolling in college courses. Students not performing well are placed in remedial courses. Test results are published annually and used to foster improved learning at the high school level. In addition, each college remedial program is evaluated annually using multiple criteria with results published. Data have demonstrated effectiveness of the State's remedial efforts.

* College Outcomes Evaluation Program. Created in 1985 by a Board of Higher Education mandate, an advisory committee of college faculty and administrators is studying how best to assess the impact of higher education in the State. Four broad areas have been delineated: (a) student learning (from basic skills to college major); (b) student development and post-college plans; (c) research, scholarship, and creative expression

(focus on faculty activities); and (d) community/society impact (focus on initiatives as a whole). A report is scheduled to be published in fall, 1987.

Questions

It is suggested that the following questions be addressed in considering how a statewide assessment program in higher education might be developed:

* How can a statewide system be developed given the diversity of student missions, programs, and institutions in the State?

* Will assessment improve higher education or harm it by diverting needed funding and energy or by focusing on narrow, overly simplistic, and misleading indicators of accountability?

* What should be the focus of assessment--individual students, programs, institutions, or the statewide system?

* What educational skills should students possess?

* Who should be tested or assessed? Who should do the assessing? When, and who should pay for it?

* How do you balance "value added" assesment with minimum competency?

* Should there be a single instrument or multiple measures? What should be assessed?

* Should a test be developed as a rising junior exam barring students from continuing unless they demonstrate appropriate competency?

* Are there other methods of assessing students' performance beyond paper and pencil tests? Are they feasible?

* What is the relationship between process and outcomes and which should be the focus?

* What data are already being collected and how should they be used?

* How do you balance access and quality?

* How do you balance accountability with goals of excellence?

* How can you balance a "top down" approach with a "bottom-up" or "grass-roots" approach?

* Can any system appropriately measure the esence of education, especially higher education, given the number and diversity of variables involved and the lack of control over these variables?

* Does college have a significant impact on students beyond normal maturation and outside influences and can this impact be determined?

Chapter 3

Summary of General Discussion

Led by Dennis P. Jones, NCHEMS

Dennis Jones: The purpose of this discussion is to try to develop some direct advice to Emerson Elliott as to what the Center can do and what the Center should do in the area of post-secondary asessment. It is my view that we must keep in mind two things: (1) the Center's traditional assessment activities such as NAEP, and (2) the Center's traditional mission of reporting on the condition of education.

If a national postsecondary assessment system were to be established, what would be its principal purpose? Should it be primarily concerned with serving national needs, or on gathering information of value to States or individual institutions? I strongly believe that the national interest should be paramount. We should endeavor to develop an indicator of the performance or effectiveness of the collective system of higher education in the United States.

Other purposes which might be served would be to facilitate comparisons: (1) among institutional sectors (e.g. liberal arts colleges vs. land grant colleges); (2) among individual institutions; (3) among different major fields; and (4) among individual students.

We might also consider the question of what levels, within postsecondary education, should be assessed. Almost everyone who has written or done research on the subject stresses the baccalaureate level and I think that we would be well advised to stick with that.

Another question to consider is, What dimensions of outcomes should be assessed? Among the possibilities are cognitive abilities, skills, values and attitudes, and a series of relationships between the individual and the institution that we measure by such indicators as retention rates, completion rates, rates of transfer to professional school, graduate school or the work force. All of these were talked about today.

John Wittstruck: Whatever the Center does should be national in scope. We should at least consider assessing the entire postsecondary domain, including vocational education, the private sector, etc.

Paul Pintrich: We can't measure everything. Why not concentrate on targets of difficulty, such as "critical thinking"?

Edward Morante: We should establish a panel of educators to meet on a reglular basis to recommend what sorts of assessments we should be doing,--to reduce duplications, etc.

Cliff Adelman: It's a mistake to bring in experts to make recommendations. All you get is generalities and cliches.

Elaine El-Khawas: It is probably impossible to characterize the implications for the Center of what has been said today. We've had an emphasis on questions, not answers. All that mirrors exactly where the discussion and debate on indicators are. That's the state of the art.

Dennis Jones: In travelling around the country I find that State level people are interested in assessing the following things:

*general knowledges and skills
*higher order thinking
*reasoning
*communications

They are convinced that these are the absolute minimum things that a college educated person should have. There is also symbolic value in assessing these things. It tells the country what we think is important.

Penny Engel: It would be desirable to tie in with some international assessment efforts so that we could make some international comparisons.

Kent Halstead: I think that the Condition of Education should give some specific examples of good assesment activities at individual institutions. This might encourage similar activities elsewhere.

Dennis Jones: Let me rough out a proposal and let you all shoot at it:

(1) We should think seriously about a postsecondary NAEP. The cohort would be new college graduates.

(2) Assessment should be of the basic dimensions noted by Aubrey Forrest, namely communication, quantitative skills, reasoning, and higher order thinking.

(3) I'd recommend staying away from individual discipline areas (major fields) with one exception: teacher education. The Center for Education Statistics would be remiss if it did not concern itself with the quality of teacher training.

(4) We should also look at indicators of values and attitudes, as in CIRP. The data are, for the most part, there and could be utilized at relatively little cost.

(5) We should attempt to assess some system-wide indicators, e.g. the percentage of a cohort that

persists through various levels of college, and the economic consequences of having or not having a college education.

David Sweet: I hear you indicating that there needs to be a process of getting to some national assessment, and that that process would require a lot of participation beyond what we have today. But perhaps today is one major start in that direction. It seems likely that any assessment we would have would not be solely of cognitive outcomes but would also take into account a variety of other differences,--differences in institutional types, differences in a variety of things we have yet to specify. We ought to take advantage of what we already have, with the IPEDS (Integrated Postsecondary Education Data System), perhaps the CIRP (Cooperative Institutional Research Program), perhaps our longitudinal studies program, and the methodology of NAEP at the elementary secondary level. We ought to build on the existing structures. We should also look at what's going on in the States and at the national level to try to find out what people already know so that we don't reinvent the wheel.

We should keep it simple at first,--something along the lines of reasoning and communication skills that Aubrey Forrest talked about. We should be careful not to assume that scores on reasoning and communication tests are identical to college scores.

Our longitudinal studies have always started at the high school level and moved on up. That is a very inefficient way to address some things in postsecondary education. We need to have our base year formulated on a postsecondary model,--not one that just happens to bump into postsecondary with a certain probability. That's about as much as I can formulate, Em. Tell me if I've gone too far.

Emerson Elliott: No, that sounds fine. Given what we've heard here today and the present state of the art, we are really at the beginning.

David Sweet: I'm absolutely delighted at the participation we've had today,--the ideas and the exchanges. We want you folks to continue as specific participants in the process.

Dennis Jones: One final comment. There are really two quite different philosophies running loose in the States, or at least two extreme ones, regarding what the States' role in assessment should be. One philosophy is that the States should do some assessments around identified state-level objectives or desirable ends which have been specified in the name of State interests. Another philosophy is implicit in the fact that many States are doing things that the institutions ought to be doing for themselves. Many of these assessments are really in the institutional domain and it's real hard to make a case that they are tied to a statewide perspective on anything.

For the immediate future I strongly recommend that CES take a national level perspective on national level concerns rather than trying to second-guess the priorities of States or the priorities of 3,300 institutions. That means doing two things: (1) being very, very clear about what's in the national interest and assessing that, and (2) being able to say it's in the national interest and that that is the reason we are assessing it. I believe that these two ideas make a lot of sense and I would recommend that you keep them in mind as you sort through what you might do.

David Sweet: I have no concluding comments but just want to thank you very much for coming and making this such a valuable day.

Chapter 4

Implications for the Center

This chapter comprises parts A and B.

Part A is the synthesis and recommendations paper prepared by Dennis P. Jones after the conference. His paper is entitled, "Postsecondary Education Assessment: The Role of the Center for Education Statistics."

Part B is comments on the conference by David A. Sweet, Director, Education Outcomes Division, Center for Education Statistics.

Chapter 4
Part A

Postsecondary Education Assessment:
The Role of the Center for Education Statistics

Dennis P. Jones
National Center for Higher Education Management Systems

The issues of whether and how to assess the consequences of our investments in postsecondary education are issues that have captured the attention of the entire higher education community. Whether one is a proponent, an opponent, or undecided, assessment is a topic that is very difficult to avoid; it is a pervasive theme that runs through almost all conversation of both political and academic leadership. Pressures to address the topic, real and imagined, are found everywhere. The responses are even more numerous than the pressures and run the gamut from outright refusal to "play the game" to direct responses to external demands to self-initiated, proactive stances.

The leadership of the Center for Education Statistics (CES) finds itself in the same position as many other actors or potential actors on the field. They are asking the questions, "What, if anything, should we do? What, if anything, can we do?" Perhaps more than other actors, they must also ask, "What are our special responsibilities?" and "What are the unique constraints within which we must work?" As part of the process of answering those questions, CES sponsored an invitational conference in Washington, D.C. on November 20, 1986. The overall objectives of the conference were to develop recommendations for the Center for Education Statistics regarding its appropriate role in postsecondary education assessment and the steps that might be taken to fulfill that role. The central questions as stated by Emerson Elliott, Director of the Center, were,

1. What is it appropriate for CES to do?

2. What is it feasible to do? (This question was raised more in a technical than an economic sense, but the economic issue cannot be ignored.)

As background information for this conference, four commissioned papers were prepared by individuals who have differing perspectives on the topic of postsecondary education born of current research and years of experience in developing and implementing assessment programs. These papers provided the basis for considerable discussion. At the end of that day, it fell to me to try to summarize the discussion, to identify the common ground, and to propose to the Center for Education Statistics a set of recommendations that could guide CES efforts regarding postsecondary education assessment. This paper

documents my attempts to fulfill this charge. In the process of writing it, it became clear that it would be necessary to step beyond the bounds of the papers presented and the conversations held at the conference to develop a coherent set of recommendations. I ask the readers to accept the paper in the spirit in which it was written--as the second step (after the conference itself) in the long process by which the Center for Education Statistics will shape for itself a specific course of action. I hope I have succeeded in organizing the issues and summarizing the areas of agreement in ways that the participants at the conference will find recognizable and that administrators and staff at the Center for Education Statistics will find helpful.

I. The Major Questions

There are many ways to approach the task that I have been assigned. I have chosen to approach it by identifying what I believe are the critical questions that must be answered on the way to resolving the broader questions of "what should be done, and what can be done." I have chosen to briefly summarize these questions at this juncture and to address each of them in the subsequent section of the paper. The major questions as I see them are as follows:

1. What is the Federal role in collecting statistics on postsecondary assessment? What is the rationale for any involvement of the Center for Education Statistics in this arena? All of the following questions become moot if it becomes clear that the Center for Education Statistics does not have a role in collecting postsecondary education information.

2. What is the unit(s) of analysis appropriate to the CES task? In other words, CES should be collecting data about the outcomes of what? The following are all potential answers to that question:

   - o A national system of higher education
   - o State systems of postsecondary education
   - o The outcomes of the education provided by different institutional sectors (research universities, comprehensive universities, baccalaureate institutions, community colleges, etc.)
   - o Individual institutions
   - o Programs or majors
   - o Individual students

3. With which levels of postsecondary education are we most concerned? Again, there are many options:

   - o Completion of a certificate program
   - o Associate degree
   - o Baccalaureate degree
   - o Post-baccalaureate degree

4. What dimensions of academic performance should be assessed? At the outset let me indicate that no attention during the conference was given to assessment of the consequences of the research and service programs of institutions of higher education. Throughout, the discussion focused on student growth and development. That focus will be maintained here. Even when one limits the discussion to assessment of the student growth and development outcomes of postsecondary education, however, one is left with a potentially very long list of potential outcomes. A useful way of categorizing this set of potential student outcomes is that described by Ewell (1984) and Jones (1985). The schema organizes student outcomes along the following dimensions:

   a. Knowledge
      - o General
      - o Specialized (discipline-related)

   b. Skills
      - o General
      - o Specialized (typically vocation-specific)

   c. Attitudes/Beliefs/Values

   d. Relationships
      - o To higher education (retention or participation in subsequent education)
      - o To the economy
      - o To professions
      - o To society at large

   Part of the task of defining any role for CES in post-secondary education assessment will be to select from this list those dimensions appropriate to the CES task.

If these four major questions can be answered fairly unambiguously, a clear direction for the CES role in gathering statistics concerning postsecondary education outcomes can be established. From this base, specific recommendations for CES action can be made. Suggested answers to these questions and the recommendations that follow are addressed in the subsequent sections of this paper.

II. Questions Addressed: Some Suggested Answers

In this section I intend to address in sequence the four questions listed above. To the extent possible, I will draw on presentations made at the conference and the free-wheeling discussions around those presentations in coming to the answers. It is true, however, that none of these questions was thoroughly resolved as a consequence of discussion at this meeting and in some cases discussion was, at best, tangential to the questions that I have posed. As a result, I have taken considerable liberty and drawn on my own thoughts in developing answers to some of these questions.

Question 1: What is the justification for the Center for Education Statistics collecting assessment data on higher education?

Throughout the day-long conference the question, "Why are we doing this?" arose numerous times. When it was answered, it was answered from an institutional perspective. In this context it was noted that the interest in assessment has increased 1) in response to demands for a different kind of accountability, 2) as a way of establishing a competitive edge in the pursuit of students and financial resources, and 3) as an element in a sincere effort to improve the quality of higher education. These are all factors that have served to ignite an interest in postsecondary education assessment on one campus or another. But what of the Federal role? Why should the Center for Education Statistics have any need to collect statistics that reflect on the performance of higher education in the United States. I find one persuasive answer and two potential answers to that question.

The one constant in the history of the Education Department from its establishment in 1867 until the present day is the charge of "collecting such statistics and facts as shall show the condition and progress of education in the several States and Territories, and of diffusing such information respecting the organization and management of

schools and school systems and methods of teaching."* In short, the Education Department has the statutory responsibility for collecting data that will allow the state of education to be monitored on an ongoing basis and for providing this information to policy-makers at all levels. While the responsibility for funding, governing, and otherwise running higher education is assigned to State and local government, there is a national underlying concern with the collection of statistics about the enterprise that has been a Federal responsibility for more than 100 years. In carrying out this mandate, the Center for Education Statistics and its predecessor organizations have, for many years, collected data about numbers of students served and resources (faculty, financial, and physical) utilized in the process. Save for information about the number of degrees granted, there has been no attention to data about outputs that balances our fascination with data about inputs. This broad, "condition of education" charge suffices to provide a rationale for CES to engage in some effort to assess the outcomes of postsecondary education. It is a general charge rather than a specific one; the response to it must be likewise general. Nonetheless, it provides a national level response to the question, "Why are we doing this?"

There is a second rationale for CES collecting assessment data, one dealing with the need to support educational policymaking at the Federal level. The Federal Government operates several hundred separate programs that impact on postsecondary education either directly or indirectly. Most of these programs have specific objectives and are (or at least could be) separately evaluated in order to determine the extent to which these programmatic objectives have been met. But what of the combined effects of these numerous efforts? Taken togther are they serving to accomplish desirable means and ends? Here again is the need for a broad description of the condition of education. To date, Federal objectives for postsecondary education have been couched largely in terms of student access and choice. If a consensus were to form around selected dimensions of student growth and development that represented a distinct Federal objective for postsecondary education, this statement would provide a specific justification for a CES role in collecting assessment data.

---

* Daniel B. Levine (Ed.), Creating a Center for Education Statistics: A Time for Action, Washington, D. C. National Academy Press. 1986, p. 5.

Given the absence of this consensus, this justification falls into the category of "potential" rather than "persuasive."

A third, and again potential, response to this question derives from the CES service role. A central data collection agency like CES can in some instances collect data more easily and more efficiently than State agencies or other potential users who have a common need for data and who would, in the absence of central agency activity, each engage in its own uncoordinated effort. This "data clearinghouse" role is well illustrated by the CES role in collecting residence and migration data--data that have relatively little Federal utility but are of interest to the States and can be collected much more efficiently and effectively through a central agency. As yet no case has been made by the States (or any other entities) indicating that they all want a particular kind of assessment done and that it would be more efficient and effective to have such assessments conducted centrally. Experience indicates that the absence of a specific request does not mean that the data would not be widely used by State governments and others if they were available.

A slightly different kind of clearinghouse role was suggested by authors of some of the papers prepared for the conference. The authors, particularly Pintrich and Forrest, emphasized the use of assessment information for purposes of evaluation and improvement at the program and campus level. In this context, they recognized a need for information that would allow institutions to assess their performance vis-a-vis national norms for similar institutions. Forrest explicitly suggests an information compilation and dissemination function "that envisions groups of cooperating colleges and universities reporting results of their data collection efforts to the Federal Government as group data... The reporting of such data to the Federal Government would be followed by dissemination of the data by the Federal Government to interested parties across the Nation." (p. 8)

In summary, the general "condition of education" requirement is likely to remain the cornerstone for any CES program in postsecondary education assessment. The potential also exists for CES to provide an important service to individual colleges and universities by disseminating data collected through a wide variety of other, original sources.

Question 2: What should be the unit of analysis?

Again there was relatively little direct discussion of this question at the planning conference called by CES. As a consequence, I will take the liberty of interjecting my own sense of the appropriate response to this question. I will argue that only the national and the institutional sector units of analysis are appropriate for CES attention at this time. It was correctly indicated by many participants at the meeting that higher education is an incredibly diverse enterprise. The finer the distinctions we try to make in the units of analysis, the more fine-grained must our assessment tools become. The converse is that the less we seek to make distinctions the more important it is to find common elements and be willing to accept those elements as common. It is not at all clear that this level of agreement can be reached. It would require that we come to grips with that basic set of knowledges and skills that we believe should be possessed by anyone who considers themselves to be educated at the college level. While it is probably far too much to ask at this time to develop standards of competence in these areas, it may be possible to develop an agreement on the "dimensions" of competence that fit within our level of expectation. This will be discussed in more detail under question 4 below.

One can make an argument for the national and institutional sector units of analysis on other kinds of pragmatic grounds. The interest of CES in a national level perspective should be self-evident. The selection of the institutional sector unit of analysis as a second perspective for CES also has a basis in pragmatism and can be arrived at by a process of elimination. Focusing on individual students or individual institutions would be neither feasible nor justifiable given the Federal role in postsecondary education. The structure of postsecondary education in the various States is so different as to make comparison across States difficult. As Clifford Adelman of OERI indicated at the conference, the development of assessment instruments for the various majors is still very much in a research mode; thus, it is not feasible for CES to develop assessment approaches at this level of analysis. This leaves the institutional sector unit of analysis as the remaining possibility. This level of analysis is attractive as the first level of disaggregation below the national system for at least three reasons. First, while admitting great variation among institutions, there is greater similiarity in purpose and mission

within a given sector than across sectors. Community colleges in two States are more likely to be similar than a community college and a research university within the same State. Second, and tied to this first point, much of the analysis in higher education is already done by institutional sector, both at CES and by a variety of informal groups and formal associations that conduct data exchange activities. For consistency, any assessment data should be organized in the same way. Finally, analysis by institutional sector is consistent with the clearinghouse role proposed for CES with regard to assessment information. To the extent that institutions look to CES for normative data of various kinds, the institutional sector unit of analysis is most appropriate.

Question 3: What levels of education should be the focus?

Without long explanations I would suggest that the associate and baccalaureate levels be the immediate target of CES attention. Baccalaureate level education has been the focal point of most of the national studies of higher education (NIE, AAC, AAHE, and soon the Carnegie report). Of equal importance is some assessment at the associate's level. Community colleges enroll large numbers of students--and disproportionate numbers of minority and lower income students that are important targets of much public policy. Post-baccalaureate education should probably be omitted from the current CES assessment agenda both because it represents a less important public topic and because, owing to its heavy disciplinary orientation, it is more difficult to assess in any general way. The exception, as is noted below, might be in those programs that prepare teachers and administrators for the Nation's school systems. Here it may by important to deal with all levels.

Question 4: What outcomes dimensions should be assessed?

In the previous section four general domains of assessment were identified. Slightly rearranged, these were

a. General knowledge and skills

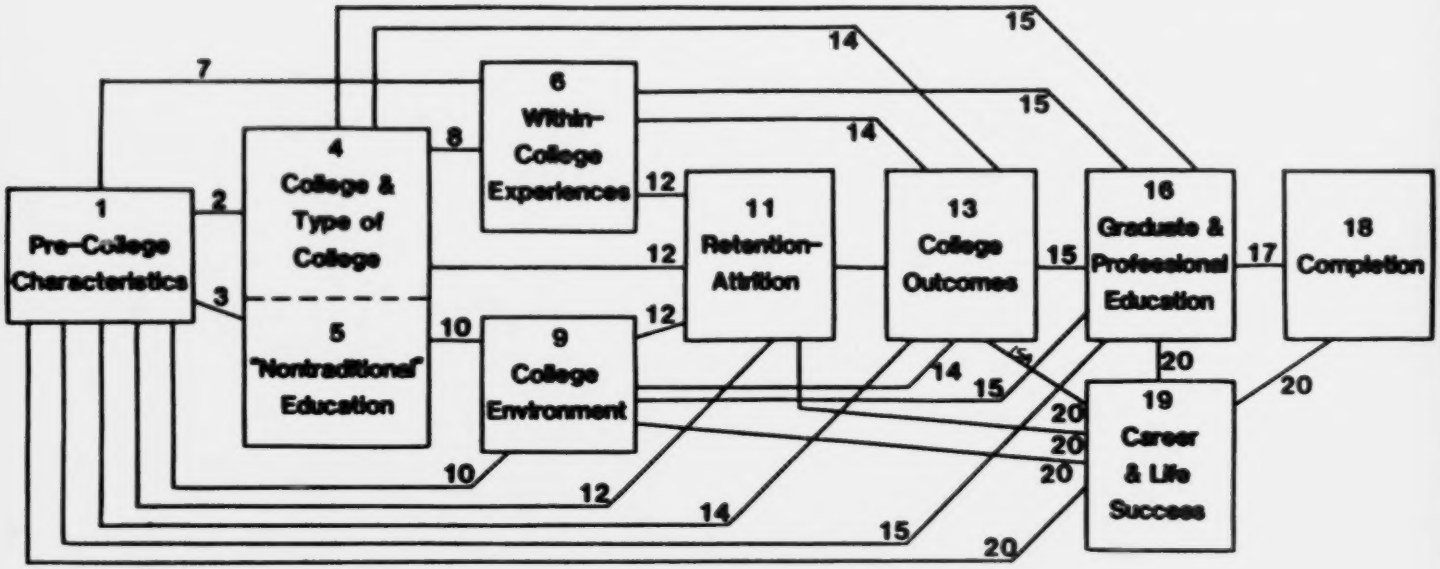
b. Specialized (disciplinary) knowledge and skills

c. Attitudes, values, beliefs

d. Relationships (to higher education, the economy, professions, etc.)

Figure 1

A MAP OF POST SECONDARY ASSESSMENT



Source: Leonard Baird, Postsecondary Assessment: The State of the Art.

Referring to the map of postsecondary education assessment provided by Leonard Baird in his paper, these dimensions (for baccalaureate and associate level education) are consistent with those identified as retention/attrition (11), college outcomes (13), the transition to graduate and professional education (15), and the transition to career and life success (the line that I have labeled 15a). In this diagram, knowledge, skills, and attitudes and values would be listed as student outcomes and retention/attrition, graduate and professional school, and career and life success correspond to the relationship dimension in Ewell's categorization schema. For ease of reference, this map is attached as Figure 1 to this paper.

Throughout the discussions at the planning conference, much was made of the real diversity in American higher education both with regard to the student body and their personal goals and the institutions of higher education and their differing missions, purposes, and priorities. Still there were two dimensions along which there was considerable agreement--it was felt that both general knowledge and skills and some of the behavioral or relationship dimensions were appropriate targets of assessment regardless of the institutional sector or student characteristics. These commonalities were identified by individuals who approached the subject from a conceptual perspective as well as those who indicated their sense of accountability requirements and political reality. From this there emerged three priorities:

1. Some assessment of the extent to which college graduates possess general knowledge and higher order skills. Much definitional work remains to be done, but this label includes not only communication and computational skills but the ability to analyze, synthesize, solve problems, think critically, and generally apply knowledge to the solution of problems. In his recitation of programs in the various states, Morante identified assessment of the basic skills of communication and computation as the dominant feature of those programs. The incentive funding program operative in Tennessee and the College Outcomes Evaluation Program being developed in New Jersey are more broadly conceived and are more consistent with assessing those dimensions of general knowledge and higher order skills normally associated with a college education.

2. Retention or persistence in the system.

3. After college experiences, particularly a) enrollment in further education and b) employment.

Specialized or disciplinary knowledge and skills were also noted as being important. However, the enormous programmatic diversity in higher education and the difficulties associated with assessing this variety of disciplines serve to relegate this category to a lower priority. I noted in my summary comments at the conference and repeat here the suggestion that CES make one exception in this area and develop a strategy for assessing the outcomes of teacher education programs on a national basis. As the statistical agency responsible for describing the condition of education, it is particularly important for CES to attempt some assessment at those points of interface between elementary/secondary and postsecondary education. Much is already being done to assess outcomes of high school graduates (or 17-year-old individuals) and thereby provide information on one of the points of interface between secondary and postsecondary education. The other major point of tangency is that represented by the graduates of teacher education programs and their entry into the work force as teachers in the elementary and secondary schools. Just as the quality (or academic preparedness) of individuals moving from elementary/secondary to postsecondary education is a concern, so is the quality of those individuals (graduates of teacher education programs) who move from postsecondary to elementary/secondary education.

The dimension that I have labeled attitudes/values/benefits received very little attention during the discussion at the conference. The orientation was heavily toward the cognitive and behavioral consequences of postsecondary education in those discussions. This is, however, an area in which many of the requisite data are already being collected and some consistent attention could yield important discriptors of the impact of college and yield important descriptors of the impact of college and yield benefits at relatively low cost. Such information could provide an additional and useful perspective on the consequences of differing higher education experiences.

In sum, I have argued that in order for the Center for Education Statistics to fulfill its mandate to report on the condition of education, the agency should be actively involved in

postsecondary education. These data should be provided from a national and institutional sector perspective, should deal with baccalaureate and associate level education, and should initially deal with the dimensions of 1) general knowledges and skills, 2) the specialized knowledges and skills required of those individuals preparing to be teachers in the elementary and secondary schools, 3) relationships of college graduates to higher education (retention and progression to higher levels of education) and to the economy (typically the fact and nature of employment).

III. Recommendations

Given this general set of directions, what are the reasonable implementation steps? This topic received relatively little attention at the conference save for that engendered by my summary comments at the end. However, since I understand it is the intent of CES to circulate this paper as well as the other commissioned papers to a broader audience for review and comment, I will suggest an agenda at this juncture as a concrete starting point for subsequent discussion.

Based on conclusions summarized above, I would suggest the following to CES:

1. Fund the development of a version of the National Assessment of Educational Progress specifically designed to assess the extent to which AA and BA degree winners have achieved mastery of general knowledge and skills (call it critical thinking, problem solving, or whatever). The young adult (age 21-25) version of NAEP that provided the basis for the recently released report on literacy is a potential model for such an assessment program. It is, however, more oriented toward basic rather than higher order skills and is not, in my mind, a substitute for the kind of assessment program that I have in mind. In the short run, however, it would be useful to include the results of that assessment in future versions of The Condition of Education inasmuch as this study does provide one clear indication of the different performance levels of individuals with different educational backgrounds.

2. An attempt should be made to develop information from a national perspective on those programs that prepare individuals for employment as elementary and secondary school teachers. If the recommendations of the Carnegie Foundation are carried through and a national certification exam is developed, a mechanism for carrying out this recommendation will be in place. It is important, however, that CES stay involved in this project sufficiently to know that the major subcategorization schemes (for example, by type of institution) can be accommodated by whatever descriptive data are associated with these tests. The current

mechanism for assessing preparedeness of teacher education program graduates is the National Teachers Exam. Its use is far from uniform across the country. It is required of all teacher education graduates in some States, mandatory for graduates of certain institutions in other States, and purely voluntary in still other States. Nevertheless, it would be worthwhile for CES to investigate the extent to which a nationally representative sample could be extracted from the set of individuals taking the NTE each year.

3. Continue to use existing sources of data to yield basic indicators of some of the relational outcomes. The National Longitudinal Study and High School and Beyond are ideal devices for calculating rates of student retention within the system of higher education in the country. Most retention studies are done on an institution-by-institution basis. However, from the national perspective it is less important that students persist in a single institution than that they persist to the completion of their educational objective. Longitudinal studies like NLS and HSB are the only existing mechanisms that allow these individuals to be consistently tracked over time. The alternative is to use cross-sectional data that show

   a. Relationships between numbers of first-time students and the numbers of degrees granted at subsequent (say, five years) point in time.

   b. Trends in proportions of specific age cohorts (e.g., 25-34) holding college degrees.

   The longitudinal approach is superior but with new cohorts being established only at 8- or 10-year-intervals there is also reason to buttress this information with cross-sectional data available from CES and the Bureau of the Census on a more frequent basis.

4. Utilize data from the Cooperative Institutional Research Program (CIRP) conducted by UCLA and ACE to more fully indicate the impact of college on students' attitudes, values, and beliefs. To be sure, this data base has its limitations, particularly since it deals only with full-time students enrolled at institutions that participate in the program. Now that a senior survey is being implemented on a regular basis, these data are such that they show changes in student values and attitudes over the course of their college careers and provide particularly useful insights as a result. There is evidence that the data can be computed in such a way that they reflect representative national samples of full-time students. The gap is in part-time students, an increasingly large proportion of the

American higher education student body. It seems that there are at least two options. First, accept the data as they are and not worry about the part-time student population in this context. The second option is to conduct an experiment to see the extent to which changes in values and attitudes of part-time students are consistent with (or different from) those of full-time students. As indicated earlier, this particular dimension of student outcomes received relatively little attention during the planning conference. However, I believe that CES would be unwise to neglect 20 years of "outcomes" data that have been systematically and conscientiously collected. They present a resource deserving of more utilization.

One final comment with regard to CIRP. The instruments used in this program include questions that might be considered a "poor man's value added" assessment--student self-reports on questions of cognitive development, etc. To be sure, these results cannot be accepted at face value. It would be worth the effort (probably through OERI) to conduct a reseach effort to determine the correlation between these self-reports and the results obtained through testing for cognitive development along similar dimensions. If the correlations were high, the self-report data would serve as an inexpensive (surrogate) method for filling in some of the time gaps in a more rigorous assessment program.

5. Take pains to ensure that high levels of community involvement are maintained as CES moves through the process of formulating and implememting its postsecondary assessment data. This recommendation was clearly enunciated by several participants at the conference and reflects significant levels of interest, commitment, and ultimately, concern within the higher education community. There is widespread recognition that increased attention to assessment is required for reasons arising both inside and outside the academy. This recognition has spurred much activity and has helped to create a good deal of interest in many quarters. In addition, however, there is the underlying reality that, in discussing assessment activities, we are necessarily discussing the purposes, goals, and objectives of American higher education. In selecting those outcomes dimensions to be the focus of assessment activities, CES is making a statement of values--these are the outcomes deemed most important from the national perspective. CES would be wise, if not obligated, to involve the broad spectrum of interests and perspectives in the higher education community as it pursues this critically important task.

References

Association of American Colleges. Integrity in the College Report to the Academic Community. Washington, DC: Association of American Colleges, 1985.

Astin, Alexander W., et al. The American Freshman: National Norms for Fall 1984. Los Angeles, Calif.: University of California, 1984.

Baird, Leonard L. "Postsecondary Assessment: The State of the Art." Paper prepared for Office of Educational Research and Improvement, U.S. Department of Education. Lexington, Ky.: University of Kentucky Educational Policy Studies and Evaluation, November 1986.

Bennett, W.J. To Reclaim a Legacy. Washington, DC: National Endowment for the Humanities, 1984.

Ewell, Peter T. The Self-Regarding Institution: Information for Excellence. Boulder, Colo.: National Center for Higher Education Management Systems [NCHEMS], 1984.

Forrest, Aubrey. "Assessment of College Outcomes." Paper prepared for Office of Educational Research and Improvement, U.S. Department of Education. Iowa City, Ia.: American College Testing Program, 1986.

Jones, Dennis P. "Indicators of the Condition of Higher Education." Paper prepared for the National Center for Education Statistics. Boulder, Colo.: National Center for Higher Education Management Systems, October 1985.

Kirsch, Irwin, S., and Ann Jungeblut. Literacy: Profiles of America's Young Adults. Princeton, N.J.: New Jersey Department of Higher Education, November 1986.

Morante, Edward A. "The State of the States in Postsecondary Assessment." Paper prepared for Office of Educational Research and Improvement, U.S. Department of Education. Trenton, N.J.: New Jersey Department of Higher Education, November 1986.

National Institute of Education, Study Group on the Conditions of Excellence in American Higher Education. Involvement in Learning: Realizing the Potential of American Higher Education. Washington, DC: U.S. Government Printing Office, 1984.

Pintrich, Paul R. "Assessing Student Progress in College: A Process-Orieted Approach to Assessment of Student Learning in Postsecondary Settings." Paper prepared for Office of Educational Reseach and Improvement, U.S. Department of Education. Ann Arbor, Mich.: National Center for Research to Improve Postsecondary Teaching and Learning (NCRIPTAL), November 1986.

Part B

Comments

David A. Sweet
Director, Education Outcomes Division

Currently there are no nationally representative data on what knowledge and skills students entering postsecondary institutions bring with them or what new skills they possess when they leave for more advanced education or training, or enter the world of work.

What we do know about the knowledge and skills of postsecondary students is very limited. It is confined to (1) students in individual institutions or selected groups of institutions and (2) selected groups of students such as those required to take an examination for admission to a particular program, school, or institution.

The practical consequence of these limitations is that we cannot generalize from these isolated pieces of information to the population of postsecondary students. Particular institutions may be informed about the education performance of their own students, but the Nation as a whole lacks an indicator of what postsecondary students know and can do.

Yet there is a growing demand for this assessment information. Students and their parents have asked for more than the present anecdotal information on the value of education and see gains in knowledge as an important part of the response they are looking for. Public policy makers and lending institutions have asked for this information as well, as have businesses faced with decisions of employment and training options.

It is in this context--on the one hand a growing interest in post-secondary assessment information, and on the other hand a substantial amount of isolated assessment activities which are very fragmented and do not lend themselves to synthesis or generalization--that this conference was initiated. The papers, presenters, and ensuing discussion, however, went far beyond merely confirming the preceived need for assessment information.

The conference provided direction for future postsecondary assessment activities at the national level. The major themes of the conference might be summarized as follows:

- CES should initiate a program to collect and report postsecondary assessment information on a nationally representative sample of students as part of its responsibility to report on the condition of education.

- The program should be developed with broad institutional input on content, procedures, and potential uses.

- The program should place a high priority on measuring a limited number of broad skill areas such as communication skills and reasoning skills. It should not attempt a detailed assessment of specific subject matter areas; postsecondary education is too broad and heterogeneous for such an undertaking.

- The program should also gather a limited amount of background information on students participating in the assessment and the institutions they attend.

- The program should provide the assessment information broken out by major categories of interest--e.g., by different institutional types and by regions of the country.

- The highest priority should be on assessment of skills at the baccalaureate degree level. Entry level skills would be in second priority.

Appendix

Postsecondary Assessment: The State of the Art

Leonard L. Baird
University of Kentucky

Perhaps the best way to begin an evaluation of the status of postsecondary assessment is with a bagful of cliches such as the following. Postsecondary education is a multibillion dollar enterprise in the United States, involving millions of people, including some of the best minds of our society. It has profound effects on the future of our economy, and, more importantly, on the nature of our civilization. In addition to its scope, postsecondary education is very complex and diverse. ·Students range from the barely literate to those with perfect scores on the SAT, from 12-year-olds to retirees, from Eskimo to inner city dwellers, and from those attending classes simply to learn about a hobby to those pursuing advanced academic or professional degrees. Colleges vary in many ways: In size, from institutions such as Deep Springs College with 20 students to those such as Ohio State which, with 55,000 students (46,000 full-time), qualifies as a small city; in selectivity, from open door colleges that accept everyone with a high school diploma or GED certificate to colleges such as Cal Tech where 99 percent of the students are from the top 10th of their high school classes; in curricula, from St. Johns, which offers one course of study to the University of Michigan with over 200 possible majors; and in student life, from those where all students are commuters to those where all live on campus. Thus, what a college "is" can vary enormously, as can the college experiences for students.

Given these cliches about the size, importance and diversity of postsecondary education, how can we make sense out of it? How can we assess such an immense and complicated social institution?

I think that we can only address those questions by carefully examining what we want to know, whether we have the conceptual tools to understand what we are concerned about, identifying the information we would need to tell us what we want to know, and by determining the extent to which it is possible and practical to obtain this information.

By far the most important consideration among these is what we want to know about postsecondary education. I have several perspectives on this question. I have recently been helping my son choose a college and will soon go through the same exercise with my daughter. As a parent, I have a number of questions that I expect are shared by other parents. Some are obvious: What are the costs, what are my son or daughter's chances of admission, what is the curriculum like, what are the requirements for degrees, what programs or facilities are available for my son or daughter's special interests? Most of these questions can be answered by the catalogue or guide books. Others become more difficult, but often can be answered from available information, such as: What are the chances a student will drop out? get A's?

go on to graduate or professional school? Finally there are questions that may be very hard to answer: What is the daily experience like? What is the intellectual climate? Are students more concerned with football or Freud? parties or Plato? Is there a sense of community among students? What happens to students like my son or daughter after going to this college? How do they grow intellectually? Do they become mature individuals? How are their ethical and social values affected? Will my son or daughter be a better person? How will he or she look back on their college years?

Besides my role as a parent, I am also a citizen-taxpayer. I have concerns about the uses of my tax dollars in my State and nationally. I am concerned about the costs, of course, but am even more concerned about the purposes or goals these dollars are put to. Are the colleges in my State meeting the current and future needs of my State and community in terms of the training they provide students? Is there provision for both excellence and equity? Nationally, I want to know the same sort of things, with some other concerns. Particularly, whether first class education is available for students with many different kinds of talents; whether able students from families of limited means are attending and graduating from colleges; whether research funds are going for the most recent trends or "hot topics" or are concentrating on fundamental issues; whether going to college makes a difference for individuals both in terms of their careers and the quality of their contribution to society; whether colleges make a difference to the economy and the culture.

Finally, as an academic and a researcher I have additional concerns. What are the implications of the rise in student careerism and concern for wealth for colleges and for students? What is the extent of "underpreparedness" among new students? What are the consequences of those facts for colleges? What is the meaning of that elusive idea, "Quality" in postsecondary education? What do students know at the end of college? Is "involvement" the way to reach excellence? What conditions promote research among faculty? How much emphasis is placed on faculty publications? What is the relationship of faculty publication activity to teaching excellence?

These various questions cover the gamut from the naive to the sophisticated, from the practical to the speculative, from the simply factual to the very interpretive. But each of them has been the object of some attempt at systematic study. That is, there have been research efforts to develop the assessment instruments needed to address these questions and various studies have used the instrument in attempts to answer the questions. These efforts have varied in sophistication and success, but the point remains that we have a considerable arsenal of instruments and information that bear on the major issues in postsecondary education. However, as I've suggested, these attempts at assessment have met with different degrees of success. Which brings me to our second and third major concern, the availability of conceptual models to help us understand the issues at question

and the identification of the information we would need to address these questions. The last concern is whether it would be feasible to obtain this information on a wide scale.

Rather than discussing these other concerns at this point let me propose a scheme--a map if you will--of major processes in postsecondary education that puts the various questions for which we want answers into focus, and which then allows us to consider the availability of models, the identification of variables, the measurement of those variables, and the feasibility of obtaining those measures on a large scale basis.

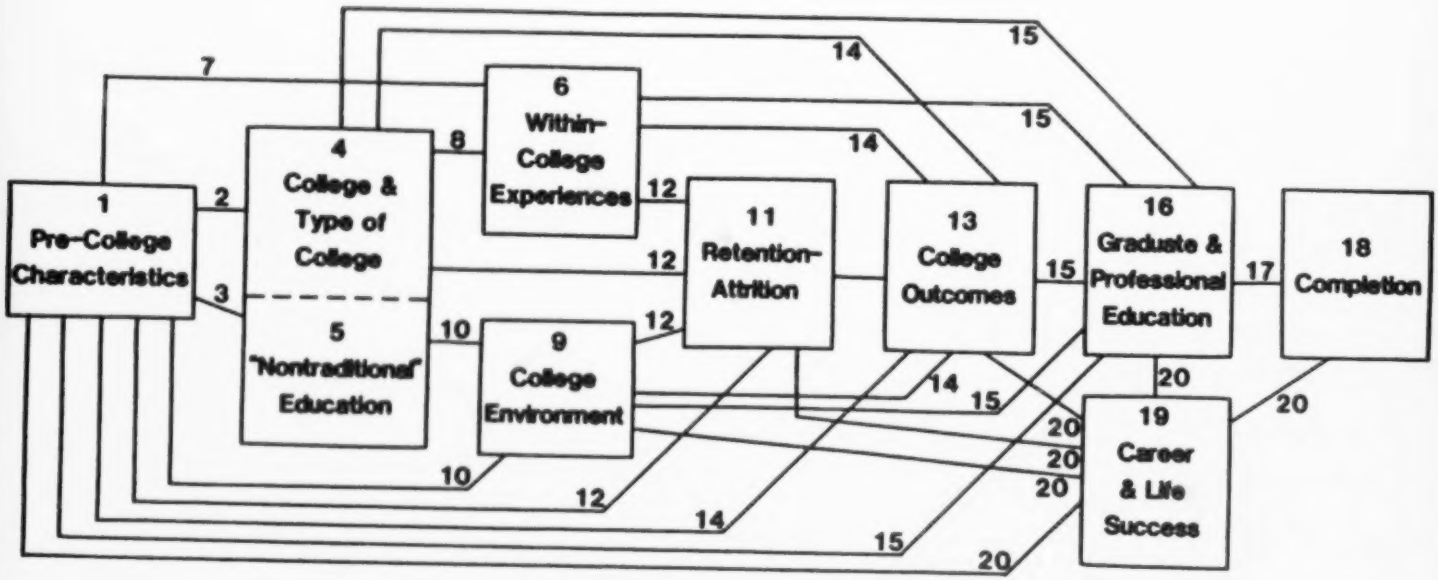
The map is shown in Figure 1. Let me define each area, and make a few comments about current issues in the area, whether they deal with conceptual models, definition of variables, measurement, or feasibility. Then I will discuss the areas that I believe would be most fruitful for further work, the importance of conceptual or theoretical models, and finally, how the entire process might be considered.

Also, let me note that this map is not meant to be a causal diagram or a totally complete description of how all the variables in postsecondary education affect one another. Rather, it is a device--a map--of how various important parts of postsecondary education are interrelated and flow into each other.

I should also point out that the map includes a great deal of information beyond what is frequently considered "postsecondary assessment" today. That is, what many people think of when they read the words "postsecondary assessment" is measurement of what students know and what their academic skills are when they apply to, or enter college, their knowledge and skills in the middle, and their knowledge and skills when they leave. These things are important of course, and involve many conceptual and technical problems. However, they are only part of the story. I think it should also be strongly emphasized that we cannot fully understand why students know what they do or do much about it until we have understood the other parts of the "map."

FIGURE 1.

A MAP OF POST SECONDARY ASSESSMENT



The first point on the map is Pre-College Characteristics. These include the level of academic preparation, educational and career goals, attitudes and views about postsecondary education, motivation, social class, age, sex, ethnicity, etc. These variables are important because they are the starting point for everything else. There are measures for virtually every characteristic. The task here is deciding which variables are most pertinent to our purposes and choosing the measure that best assesses the variable. (See, for example, the discussion of academic preparation in such sources as the American College Testing Program, [1976], the role of a range of personal factors in college admission in Willingham and Breland, [1982], and the roll of social class and cultural sophistication in preparation for the subtleties of college life, Feldman and Newcomb, [1969]). Point number 2 on the map is the high school-college transition, which is concerned with how students choose to attend college, the influence of finances, access, gender, social class, ability, ethnicity, etc. on college attendance. This area has been the subject of a great deal of research. I think the task here is choosing among various explanatory models and philosophical interpretations (e.g. see Manski and Wise 1983; Zemsky and Oedel, 1983; Lowery, 1982). In contrast, Number 3, adult entrance to postsecondary education is not nearly as well understood but is put into this map because it is becoming an increasingly important social fact (Peterson, 1979; Cross, 1981; Cross and McCartan, 1985). I will discuss this area in more detail later.

The fourth area--colleges and college characteristics--concerns our understanding and assessment of the important distinctions among colleges, and their influences on the flow of students to different postsecondary options. We know that research universities differ in many ways from denominational colleges. We also know that the students who attend community colleges have a different aggregate profile from those who attend selective liberal arts colleges. The challenge here is to interpret the significance of these differences. (The series of volumes on different types of colleges prepared for the Carnegie Commission on Higher Education, although slightly dated, provide many facts and insights into the significance of these differences. Some examples are Astin and Lee's (1972) portrait of the largest group of institutions, small private colleges with limited resources; Dunham's (1969) profile of State colleges and regional universities, and Greeley's (1969) description of Catholic colleges. Other, more recent portraits include Cohen and Brawer's (1982) account of the community college, and Fleming's (1984) portrait of black colleges.) However, perhaps a more important concern is the nature of various "nontraditional" forms of post-secondary education and the flow of people into them (Point 5 on the map). It has been estimated that the majority of postsecondary educational instruction is conducted in such non-traditional settings as corporations, organizations, governmental agencies, and community groups. It is very difficult to assess this tremendous diversity of educational experiences, but it is probably true that many of them represent high level instruction and learning, however brief they may be.

I think it is critically important that we understand the scope of these activities, assess the quality of the instruction, and see how the educational outcomes of such experiences can translate into the credential requirements of traditional forms of postsecondary education (e.g. see Keeton, 1980; Knapp, 1981; Scott, 1985). The sixth point on the map involves the assessment of the types of within-college experiences. That is, we know that some of the most important effects upon students during their college years are produced by choice of major and residence grouping, etc. The point here is whether we have the proper characterizations of the collegiate experience. For example, there is convincing evidence that living on campus or commuting can have substantial effects on students' collegiate careers. However, do we have any ideas to explain why living on campus or commuting have their effects that go beyond common sense? (e.g. see Chickering, 1974; Pascarella, 1985). Perhaps the most important of the choices students make within college is the choice of major, because that choice bears directly upon students' educational experiences and careers (Holland, 1985). The seventh point in the model is the influence of pre-collegiate characteristics upon within-college experiences. The importance of this point is underlined by the fact that fewer and fewer students with high test scores are choosing to major in primary or secondary education, leading some, such as the Carnegie Foundation, to speculate that we may not have enough capable school teachers in the future. There is a considerable literature that shows that students choose majors based on their backgrounds, abilities, interests, and their perceptions of the job market (Holland, 1985). The evidence on other choices is less substantial, but it is clear that students choose experiences consistent with their characteristics, and that understanding this process offers one of the main fulcrums by which policy can affect students (See Weidman, [1984] for some evidence). For example, scholarships for students considering a career in school teaching may lead them to follow through on that choice.

The next point on the map (8) is the influence of college characteristics on these choices of within-college experiences. An example of this is the evidence that students tend to drop out even more than expected from their characteristics when they attend 2-year colleges and less than expected when they attend residential liberal arts colleges largely because the within-college experiences differ. Here, however, the conceptual problem may be that we lack general theoretical models of how colleges affect students' choices. However, (See Pascarella's [1985b] general causal model).

Point 9 on the map is the college environment, which is the subjective nature of the college experience. Some of the major dimensions of the environment, identified by a variety of methods, are the sense of community, the degree of academic rigor, and the level of formality (Baird, 1980; Moos, 1979). In the last several years much more attention has been devoted to the environment, particularly to how it leads to "involvement."

Point 10 is the influence of types of college and student characteristics on the environment. An old debate concerning the environment is whether it is aggregate student characteristics that make the environment or whether the environment is created by something external to the students. That is, for example, whether it is the presence of many able students that create a sense of academic rigor or whether it is the standards and demands of the college that do so (e.g., see Feldman, 1971). There are a variety of methods to assess the environment, but little agreement on the best way to understand the environment (Baird and Hartnett, 1980).

The next point on the map (11) is simply the facts on retention and attrition. There have been numerous attempts to define and codify the possible meanings of retention and attrition. These can vary greatly (Pat Terenzini, a major researcher in this area said that he can give over 50 different responses to the question "what is the dropout rate?", all of which are factually accurate.) However, there is some consensus on definitions and it has been possible to chart the extent of retention across types of colleges, for students with different characteristics, and over time (Tinto, 1982, Noel, 1985).

A related area, (12) on the map, is the prediction or understanding of the retention/attrition process. This is one area where there are testable conceptual models which have been the subject of considerable research. These have led to new assessments of theoretically important variables. I will expand on this point later.

The next point on the map (13) concerns college outcomes, which Aubrey Forrest discusses elsewhere in this monograph. About 10 years ago NCHEMS had an extensive project to define and measure these outcomes, producing, among other documents, A Structure for the Outcomes of Postsecondary Education, (Lenning et al., 1976). That volune listed 10 categories of characteristics, such as "Competence and Skills" and over 50 somewhat more specific areas such as "intellectual skills," which of course has many sub-elements. The point is that there are many possible outcomes of higher education, including virtually every human characteristic. Clearly, the task here is what to focus on, deciding what is most important to consider. A very reasonable approach is that of Bowen (1977) who attempted to describe a consensus about what the goals of postsecondary education are, and relating the assessment of outcomes to these goals. Clearly, the choice of outcomes depends on one's values and interpretations of the purposes of postsecondary education. The appropriateness and technical quality of possible assessments depends on the choice of outcomes. The situation is complicated by the fact that many observers argue that the pluralism of postsecondary education requires each institution to have its own set of goals and outcomes. Here, perhaps more than any area, the issue is the logic of the choice we make in choosing which outcomes to study. (For general discussions of outcomes, see Lenning [1976] and Ewell [1985].)

The next point on the map (14), college effects, concerns the general influence of colleges and their programs on student outcomes. As the discussion of the last point would suggest, consideration of the variety of outcomes produces a complicated picture. However, virtually all of the research on college effects deals with change or gains on a relatively small group of outcomes: career choices, educational aspirations, and academic achievement tests. The emphasis here should be on the words change and gain. Essentially, college effects research is concerned with the differential impact of colleges, i.e., why one college has a more positive influence than another. For example, once you control for the ability and background of students, does Harvard have any better effects on students than Mississippi State? (Note that this is a different--and more sophisticated--question than the more simple minded question of the "value added" or "talent-development" approaches, which focus on single colleges.) The point is to attempt to attribute change on growth in student characteristics to the college characteristics or environment, controlling for the students' initial status. This creates many problems, since students' final status is highly determined by initial status. The assessments in this area are subject to a wide variety of logical and psychometric considerations. These include the usual concerns with reliability and validity in their multiple meanings. But they also involve considerations of the sensitivity of the measures to real change, and the meaning of the measures at the beginning and end of postsecondary education (for example, a career choice of professor or physician may be a vague aspiration for a freshman, but may be based on a much more realistic self-evaluation for a senior.) This area is fraught with problems of logic, measurement, statistical design and evidence (see Pascarella, 1985, for a trenchant discussion of these points.) In sum, there is great sophistication in this area, and a high level of understanding, but the evidence to date shows few consistent or powerful college effects. This may be due to the lack of the most appropriate measures, or to the relatively small impact of any new educational experience on students who have had 12 previous years of study.

The next point on the map (15), the transition to graduate or professional education, has not been well studied, largely because of the logistical problems of conducting longitudinal studies with college graduates. (There are some exceptions, such as Baird, 1976, Ethington and Smart, 1986.) However, many of the same variables that influence the high school-college transition also influence the transition to graduate or professional school, such as previous academic performance, the requirements of one's career field, etc. It is difficult to summarize all of these variables, and the theories of career choice and educational aspirations are often not helpful. For example, in recent years, large numbers of the students who have chosen to pursue MBA degrees have little intrinsic interest in business, but are simply reacting to their perceptions of the job market (more on this later). In sum, identification of and assessment of the important variables in this area seems to be complicated and

incomplete, especially for the large numbers of older students continuing their educations. For example, what do the GRE scores of a 35-year-old applicant, who has been away from institutionalized education for 14 years, mean?

The next point on the map, (16) the assessment of the types, characteristics and environments of graduate and professional education, has seldom been studied systematically. Although a great deal has been written about the professions and the process of professionalization, there have been few empirical studies comparing advanced education across disciplines, and even fewer studying differences within a discipline, e.g., how the environments for learning differ across medical schools. In addition, the existing work has focused almost entirely on the more prestigious professional schools such as law and medicine, or doctoral study at the elite graduate departments in traditional letters and science fields (e.g. Baird, 1974, Clark, Hartnett and Baird, 1976, Katz and Hartnett, 1976.) Very little has been done in the less prestigious professional fields or at the master's level, which is where the largest share of the enrollment is. However, the existing research suggests that professional and academic disciplines differ widely among each other and within the disciplines. I think this is a very promising area for the development of models and measures.

The next point on the map (17), concerns attrition and retention in advanced studies. (Lines linking this to earlier variables are not shown to simplify the diagram.) Partly because of the difficulties in tracking students in often highly individualized programs, there is little research in this area. For example, is an ABD a dropout? Is a student who has spent 10 years in studies without obtaining a doctorate making normal progress? If not, as many as a quarter of the graduate students in some disciplines are not making normal progress. Despite the logistical problems involved, this area is very important to understand, and one that would profit from even the simplest of studies.

The next point on the map (18), completion status, is an area where it is very difficult to know what we want to understand. For example, we might like to know how much the recipients of various degrees have learned. Although a few professions such as law have external examinations required for final admittance to the profession, for most disciplines in graduate and professional education it is unclear as to what we would look for, although there have occasionally been reviews of dissertations by external evaluators. Although much has been written about academic socialization and professionalization, there are no clear criteria by which these might be assessed. This area seems to lack both models and measures.

The next point on the map, (19), career or life success, is considered by some to be the most important area of all (Lenning et al., 1975). However, as I have written elsewhere (Baird, 1985) the assessment of "success" is quite problematical. For

example, the clearest kinds of criteria of "success" apply only to a few--such as publications and citations among Ph.D. recipients who work in academe. Most careers involve complex and multiple indicators of "success." Even such seemingly objective criteria as annual salary are very problematical. And such complex careers as medicine can have a bewildering number of possible criteria, many of which are negatively related. (For example, the most thorough attempt to define success in the physician's role resulted in some 80 measures, many of which were negatively related [Price et al. [1973].) The final point on the map (20) is the prediction of career and life success, an area that has been the subject of considerable debate, involving many political and philosophical questions, which I will not go into here. I will just note that, despite my comments on success within certain areas, and its problematical nature, at a very gross level it is possible to roughly assess general "success" in terms of educational and occupational attainment. The sociological literature is full of models attempting to account for these variables in American life (Baird, 1985).

So, having described the "map," where does it leave us? I think there are several implications from our consideration of the map. One is the content, one is the use of models, and one is our interpretation. On content, I would like to suggest that there are several areas in which we need to improve our understanding and our assessments. One area concerns the increasing numbers of adults who enter postsecondary education for the first time, or who are returning to pursue further education. It is unclear as to what methods are appropriate to assess their readiness for college or graduate education. Although older applicants as a group score lower on admissions tests, they often do much better than predicted in their classes, so some other variables are operating. But what are they? One possibility derives from various conceptions of the growth of intelligence which suggest that its meaning and form change over the life span (Schaie and Parr, 1981; Berg and Sternberg, 1985). What are the appropriate ways of assessing and matching instruction and colleges to these changing abilities? In general, although there are many small studies of adult learners in higher education, we would greatly benefit by some large scale studies of their characteristics, motivations, learning styles, and achievement.

A second large gap is information about graduate and professional education. Although we know about the total enrollment and the numbers of degrees awarded in different disciplines, we could profit from much more information. Although there are a handful of studies of the factors influencing attendance in graduate and professional school, attrition in graduate school, and the graduate or professional school experience, there is a crying need for more information in this area. As I noted before, most of what is known is based on elite professional and graduate schools and misses the experiences of the great majority of advanced students.

Yet another gap in our information concerns the attainments, plans, aspirations and views of college seniors. Here, I am not concerned with assessing outcomes in the ways that most researchers in outcomes are concerned with. I am talking about the kinds of information collected by the ACE/UCLA Cooperative Institutional Research Program in its freshman surveys. This survey includes data on students' career choices, educational plans, financial indebtedness, attitudes toward education, views on social issues, reports on their academic performance and experiences, etc. If this same kind of information were routinely obtained for seniors, it could allow us to chart trends in such areas as student career choices, academic performance, views of the purposes of college education, plans for further study, academic and social experiences in college, students' satisfaction with their college, etc. We could also compare these data for students in different kinds of colleges, in different majors, and for different groups of students, such as minority and majority students, women and men. And, if these data were accumulated over a number of years, as the CIRP data have been, we could trace a number of variables that are important in postsecondary education. One example is tracing grades received in college to examine grade inflation, i.e., to see whether an A average is more or less common from year to year. Another example is data on student indebtedness. It is often contended that with rising college costs and declining financial aid, many students are ending college owing large amounts to either the government, their colleges or lending institutions. With the kind of routine data collection I am referring to, we could chart the level of indebtedness and, at least indirectly, relate it to policy. Still another example is student career choices (not major) which some contend have become much more oriented toward financial success, leading students to plan careers as business executives, accountants, engineers, etc. and to avoid careers in such low paying areas as school teaching, social work, and public service. A senior survey would allow us to evaluate the accuracy of such claims. In addition, if it were possible to link the responses from the CIRP freshman survey to the senior survey, it would be possible to conduct studies of how students change during the course of college, and how various experiences influence these changes. There are some methodological problems in this area, such as dropouts, correctly controlling for initial status or characteristics, etc., but I think the value of the information gained makes dealing with such problems worth the effort.

The information collected would be at least as useful to colleges as the CIRP freshmen data, since students could be asked about their reactions to their college's programs and services. In these times, when colleges are being called upon to demonstrate that their graduates have gained from their programs, and to show that students are satisfied with the quality of their education, this kind of data would seem to offer a great deal of value to colleges responding to such calls.

In addition to proposing a senior survey, I would like to recommend that data from the American College Testing Program and the Educational Testing Service--the two major testing programs for the high school-college transition--be analyzed to yield data that would meet some of our national concerns. They could routinely provide profiles for students with different characteristics. For example, by routinely breaking down results by ethnic group they could provide considerable data about American minority students who are bound for college. Note here, that I am not calling for comparisons of test scores so much as for descriptions of the goals, interests, high school accomplishments and plans of students from various groups. I'm sure the reader can think of other possible uses for these vast data bases. However, the main point is to capitalize on these sources in useful and imaginative ways. Certainly these data can be used for many research studies that also bear on educational issues. For example, I did a study of the consequences of attending a large or small high school for students' activities and accomplishments, based on the "behavior setting" theories of Roger Barker, using ACT national data (Baird, 1969.) These analyses showed, for a large national sample, that Barker was right. The rates of student participation, involvement, and leadership were considerably higher in small high schools than in large high schools. Thus, for the attainment of many educational goals, it is probably better to keep small high schools rather than to consolidate them. This interpretation has many possible implications for school districting decisions. The point is that it was based on routinely collected information that would have been very costly and difficult to obtain otherwise.

An example of a possible analysis using ACT data would exploit the vocational interest test that they have administered for several years. Specifically we could compare the measured vocational interests of students who currently say they are going to major in business and other fields with the interests of students who said they were going to major in those fields in the national data of some years ago. This would suggest whether students are making choices less consistent with their interests, perhaps due to their perceptions of which fields are marketable. Many other questions could be addressed using this data. Since the data of the national testing programs has already been collected and processed, and is based upon immensely large samples, it would seem to be a resource--and a very inexpensive one--that could be used to address many questions.

Note that to this point, I have said nothing about the technical side of assessment, i.e., new psychometric approaches, the possibilities of computerized assessments, sampling procedures, statistical models, etc. Nor have I attempted to review specific measures and their strengths and weaknesses, although there are many intriguing recent developments, such as various measures of students' personal and moral maturity and Robert Sternberg's attempts to assess cognitive capabilities based on recent models of how the mind functions. Although useful, I don't think the advances we have made in postsecondary

assessment are due to such technical improvements. Rather I have concentrated on the major questions that I think we want to answer because I believe true progress in postsecondary assessment comes from developing an understanding of the areas we are concerned with, and the construction of testable models. That is, I believe we have the technical tools to develop assessments of most of what we are interested in studying. The task is to develop concepts and models. Let me give an example.

There have been hundreds of studies of attrition in higher education for at least 50 years. For many years these studies were entirely empirical, searching for some measures that would lead to better prediction of student attrition and retention. Study was piled upon study with no advances in our understanding or prediction of attrition. Then in 1970 Spady proposed a model of attrition-retention, which was adapted by Tinto in 1975. Instead of a shotgun approach these researchers proposed that students enter college with varying degrees of "goal commitment," or the value they place on graduating, and "institutional commitment," or the value they place on the particular institution they are attending, as well as their academic preparation and backgrounds. Interacting with the academic and social systems of their college, students have various experiences which affect the extent to which they are integrated into the social and academic life of the college. The level of academic and social integration then affects their goal and institutional commitment during the course of college. When students are integrated and have a high level of goal and institutional commitment they stay in college; when they are not well integrated and their goal and institutional commitment is low, they leave. Thus the model stipulates how different student characteristics and college experiences interact to affect the decision to stay or drop out. This model has been tested in a wide variety of studies. Not every prediction from the model has been supported consistently, but it has increased our understanding of the processes involved in attrition and it has led to the search for better assessments of the variables in the model. The search for better assessments in the model has led, in turn, to reconsideration of other ideas. For example, "academic integration" has many possible elements which revolve around how a student begins to feel part of the academic life of a college. One obvious element is interaction with faculty outside of class. However, there are several possible kinds of faculty-student interaction. Some analyses (Pascarella and Terenzini, 1978) suggest that the most important kinds of interaction are those that focus on academic advising and discussions of campus issues--not discussions of personal problems, general issues, etc. Thus, the attrition-retention model has led to both attempts to produce better assessment of the variables in question and to a more thorough understanding of the nature of students' college experiences. The Spady-Tinto model is an example of a model that has been developed and tested.

Another area that I believe will be extremely fruitful in the future is one where a model seems to be developing: the assessment of the meaning of "involvement." Although Astin (1985) has some ideas he labels a "theory of involvement," there still needs to be a tighter set of concepts and clear specification of how and why the elements in the theory affect each other. However, there has been at least one important attempt to assess the extent and significance of student "involvement." This is Pace's (1984) College Experiences Questionnaire, which is designed to estimate a student's quality of effort in various areas of college life. Analyses using this instrument have indicated that effort in certain areas promotes progress toward goals in those areas. Almost certainly there will be other attempts to define the meaning and assess the components of involvement that will increase our understanding of the interaction of student and college. The area of "involvement" seems to be an example of what may be gained from a model in the making.

Finally, I would like to turn to our interpretation of assessment information and how it can increase our understanding of postsecondary education. In this case various kinds of information about how postsecondary education is changing in response to social changes, can lead to ideas about the nature of those changes, which then lead to further considerations of postsecondary education. For example, as I've mentioned, numerous observers--most recently Carnegie--have pointed to indications that students have become increasingly careerist in their choices and orientations over the last 20 years (Katchadourian and Boli, 1985). This has happened almost simultaneously with the increase in educational opportunity. These facts, combined with evidence that the economic return on education has declined, have led to reconsiderations of the meaning of postsecondary education in American society (Collins, 1979). In a phrase, this has led to the idea that postsecondary education has changed from opportunities to ultimatums. Let me briefly outline this argument. Before and after World War II a college education was an opportunity for people who wished to move up in American society. That is, admission to college was a privilege, and the completion of college was almost a guarantee of a well paying and satisfying career. It was not always important what one's degree was in, as much as it was that one had a college degree. Naturally, individuals and public policy makers looked to increased educational opportunity as a way to increase the life and career opportunities of many segments of our society. There were movements for open admissions, large scale financial aid, and majors designed to meet the needs of students. And these policies seemed to work. Many more students attended college, many more graduated, and it seemed that the egalitarian goals of the policies had been successful. There were opportunities for most, if not all. However, what seemed to happen is that the meaning of a college degree began to change. Instead of a rarity it became relatively common. It was no longer an entree into a wide variety of careers. Since there were so many college graduates, employers began to look for graduates with degrees in

just the fields they were interested in, and, since a degree was not necessarily a guarantee of talent, they began to pay more attention to where graduates had obtained their degrees. Thus, one of the unexpected consequences of the success of the egalitarian reforms in postsecondary education was to make prestige and specific training more important rather than less. However, an even more important point is that a degree, once seen as almost always leading to success was now seen simply as another requirement to get into the game, i.e., an ultimatum. One solution to getting more out of one's degree was to choose fields that promised success, e.g., business, engineering, etc., which results in the "careerist" tendencies we have seen among students. Another was to go to a more prestigious college, which is reflected in the desires for status that the Carnegie studies have noted. A third was to up the ante and obtain more degrees, which is reflected in the rise in graduate and professional school enrollments. I do not wish to argue whether this conception is correct. I am putting it forward as an example of the use of general models, which allow us to make sense of the overall changes in the postsecondary system. This is also an example of what one might call a metamodel, that is, one that steps above the level of specific domains and attempts to put the entire process of postsecondary education into an understandable picture. It suggests that some other measures might be needed, such as students' views of the economic payoffs of various majors, schools, and degrees, and the extent to which their perceptions of reality affect their choices. But most of all it helps our perceptions of what is really happening in higher education, and what our "facts" signify.

I have covered a wide variety of topics in this article, attempting to focus on what we want to know, and the gaps in our knowledge. I have emphasized formulating the right questions and the use of models because I believe that the "state of the art" in assessment in postsecondary education today is not due to technical advances, but to increases in our understanding. What is important is not so much the quality of our methods as the quality of our ideas.

References

American College Testing Program (1976). Assessing Students on the Way to College. Iowa City: American College Testing Program.

Astin, A.W. (1985). Achieving Educational Excellence. San Francisco: Jossey-Bass.

Astin, A. and Lee, C.B.T. (1972). The Invisible Colleges. New York: McGraw-Hill.

Baird, L.L. (1969). "Big School, Small School: A Critical Examination of the Hypothesis." Journal of Educational Psychology, 60, 253-260.

Baird, L.L. (1974). Careers and Curricula. Princeton, N.J.: Educational Testing Service.

Baird, L.L. (1976). "Who Goes to Graduate School and How They Get There." In J. Katz and R.T. Hartnett (Eds.) Scholars in the Making. Cambridge, MA: Ballinger.

Baird, L.L. (1980). "Importance of Surveying Student and Faculty Views." In L.L. Baird and R.T. Hartnett and Associates. Understanding Student and Faculty Life. San Francisco: Jossey-Bass.

Baird, L.L. (1985). "Do Grades and Tests Predict Adult Accomplishment?" Research in Higher Education 1985, 23, 1-85.

Baird, L.L. and Hartnett, R.T. (1980). "Directory of Lending Instruments for Assessing Campus Environments." In L.L. Baird and R.T. Hartnett. Understanding Student and Faculty Life. San Francisco: Jossey-Bass.

Berg, C.A. and Sternberg, R.J. (1985). "A Triarchic Theory of Intellectual Development in Adulthood." Developmental Review, 5, 334-370.

Chickering, A. (1975). Commuting Versus Resident Students. San Francisco: Jossey-Bass.

Clark, M.J., Hartnett, R.T. and Baird L.L. (1978). Assessing Dimensions of Quality in Doctoral Education: Summary of a Multidimensional Approach. Washington D.C.: Council of Graduate Schools.

Cohen, A.M. and Brawer, F.B. (1982). The American Community College. San Francisco: Jossey-Bass Publishers.

Collins, R. (1979). The Credential Society: An Historical Sociology of Education and Stratification. New York: Academic Press.

Cross, K.P. (1981). Adults as Learners. San Francisco: Jossey-Bass Publishers.

Cross, K.P. and McCartan, A. (1984). Adult Learning: State Policies and Institutional Practices. Washington, D.C.: Association for the Study of Higher Education.

Dunham, E.A. (1969). Colleges of the Forgotten Americans: A Profile of State Colleges and Universities. New York: McGraw-Hill.

Ethington, C.A. and Smart, J.C. (1985). "Persistence to Graduate Education." Research in Higher Education, 24, 287-303.

Ewell, P.T. (1985) (Ed). Assessing Educational Outcomes. San Francisco: Jossey-Bass.

Feldman, K.A. (1971). "Measuring College Environments: Some Uses of Path Analysis." American Educational Research Journal, 8, 51-70.

Feldman, K.A. and Newcomb, T.M. (1969). The Impact of College on Students. San Francisco: Jossey-Bass Publishers.

Fleming, J. (1984). Blacks in College. San Francisco: Jossey-Bass.

Greeley, A.M. (1969). From Back Water to Main Stream: A Profile of Catholic Higher Education. New York: McGraw-Hill.

Holland, J.L. (1985). Making Vocational Choices: A Theory of Vocational Personalities, and Work Environments (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Katchadourian, H. A. and Bolin. (1985). Careerism and Intellectualism among College Students. San Francisco: Jossey-Bass.

Katz, J. and Hartnett, R.T. (Eds.). (1976). Scholars in the Making. Cambridge, MA: Ballinger.

Keeton, M.T. (Ed.). (1980). Defining and Assessing Quality in Experiential Learning. New Directions in Experiential Learning, No. 9. San Francisco: Jossey Bass.

Knapp, J.E. (Ed.). (1981). Financing and Implementing Prior Learning Assessment. New Direction for Experiential Learning, No. 14. San Francisco: Jossey-Bass.

Lenning, O.T. (Ed.). (1976). Improving Educational Outcomes. San Francisco: Jossey-Bass.

Lenning, O.T. and others. (1975). The Many Faces of College Success and Their Non-intellective Correlates. Iowa City: American College Testing Program.

Lowery, W.R. and Associates. (1982). College Admissions Counseling. San Francisco: Jossey-Bass.

Manski, C.F. and Wise, D.A. (1983). College Choice in America. Cambridge, MA: Harvard University Press.

Mois, R. (1979). Evaluating Educational Environments. San Francisco: Jossey-Bass.

Noel, L. (Ed.). (1985). Improving Student Retention. San Francisco: Jossey-Bass.

Pace, C.R. (1984). Measuring the Quality of College Student Experiences. Higher Education Research Institute, UCLA, Los Angeles.

Pascarella, E.T. (1985a). "The Influence of On-campus Living Versus Commuting to College on Intellectual and Interpersonal Self-concept. Journal of College Student Personnel, 26, 292-297.

Pascarella, E.T. (1985b). "College Environmental Influences on Learning and Cognitive Development: A Critical Review and Synthesis. In J.C. Smart (Ed.) Higher Education: Handbook of Theory and Research, Vol. 1. New York: Agathon Press.

Pascarella, E. and Terenzini, P. (1978). "Student-faculty Informal Relationships and Freshman-year Educational Outcomes." Journal of Educational Research, 71, 183-189.

Peterson, R.E. and Associates (1979). Lifelong Learning in America. San Francisco: Jossey-Bass.

Price, P.B., and Others. (1973). Measures and Predictors of Physician Performance: Two Decades of Intermittently Sustained Research. (Dept. of H.E.W. Grant No. PM00017 Report). Salt Lake City: University of Utah.

Schaie, K.W. and Parr, J. (1981). "Intelligence." In A.W. Chickering and Associates. The Modern American College. San Francisco: Jossey-Bass.

Scott, J.A. (1985). "Integration of Non-traditional Programs into the Mainstream of Academic Life." Innovative Higher Education, 9, 81-91.

Spady, W. (1970). "Dropouts from Higher Education. An Interdisciplinary Review and Synthesis." Interchange, 1, 64-85.

Tinto, V. (1975). "Dropout from Higher Education: A Theoretical Synthesis of Recent Research." Review of Educational Research, 45, 89-125.

Tinto, V. (1982). "Limits of Theory and Practice in Student Attrition." Journal of Higher Education, 53, 587-600.

Weidman, J.C. (1984). "Impacts of Campus Experiences and Parental Socialization on Undergraduates' Career Choices." Research in Higher Education, 20, 445-474.

Willingham, W.W. and Breland, H.M. (1982). Personal Qualities and College Admissions. New York: College Entrance Examination Board.

Zemsky, Robert and Oedel, Penny (1983). The Structure of College Choice. New York: College Entrance Examination Board.

Assessing Student Progress in College:
A Process-Oriented Approach to Assessment of
Student Learning in Postsecondary Settings

Paul R. Pintrich

National Center for Research to Improve Postsecondary
Teaching and Learning (NCRIPTAL)

and

The School of Education
The University of Michigan
Ann Arbor, Michigan

The role of assessment in education has a long history in educational research and policy (Haney, 1984). As Linn (1986a) points out, discussions of testing and assessment issues are ubiquitous in American education. Moreover, both Haney (1984) and Linn (1986a) note that past and current calls for educational reform have used test results to buttress arguments about the need for improvement in the educational system as well as suggested that assessment can be one tool to facilitate change. In general, the calls for reform have been aimed at improving our elementary and secondary schools, but recently there has been increased interest and emphasis on improving postsecondary educational institutions (e.g., Boyer, forthcoming; Mortimer et al., 1984). Paralleling the elementary and secondary reports, the reports on postsecondary education have suggested that assessment can play an important role in improving our efforts at the college level.

There are, of course, a variety of issues to consider in discussing assessment issues at the postsecondary level. Recently, Alexander and Stark (1986) have suggested eight parameters that should be considered in any discussion of assessment. Although this paper will focus on two of the most crucial parameters, all eight are summarized briefly in order to provide a context for the issues raised in this paper.

Alexander and Stark's Eight Parameters of Assessment

1) What are the purposes or incentives for assessment? For example, assessment can be used to certify students' credentials, identify special needs of students (both remedial and accelerated), and assist the educational institution in selecting the proper placement of the student to meet these special needs. Assessment activities can be undertaken to satisfy State or Federal reporting requirements and may differ from institutional goals for assessment. Assessment also can be directed towards program improvement or program evaluation purposes. Assessment can be focused on improving teaching and

learning as well as assessing the worth and quality of the outcomes of a program or the value added by a program (Adelman, 1986; Alexander and Stark, 1986; Hartle, 1986; Mortimer et al., 1984; Linn, 1986a; National Commission on Excellence in Education, 1983).

2) What is the nature and type of assessment data that will be collected as part of the assessment program? For example, is information going to be collected on student academic performance over the course of college (e.g., Loacker, Cromwell, & O'Brien, 1986), or on student motivation, attitudes and values (e.g., Astin, 1977), or on student employment rates after college? Will the data tap a variety of domains of interest to avoid mono-operation bias (Cook & Campbell, 1979)?

3) What is the level of analysis that will be used to guide data collection and interpretation in the assessment program? Will the data be collected at the individual student level and then aggregated to a departmental or institutional level? Will the data only be collected at the departmental, program, college, or institutional level?

4) What is the form of the assessment process? Will multiple measures be used to collect data? Will different methodologies be used to avoid mono-method bias (Cook & Campbell, 1979)? Will norm-referenced or criterion-referenced standards be used to interpret the data?

5) What agency or body will be responsible for the administration of the program? The program can be run by external parties including Federal, State, and local agencies as well as internal (to the college) institutional-wide centers for assessment, departmental faculty groups, or even individual faculty members.

6) What will be the locus of evaluation for the assessment data? For example, who will make the decisions and evaluative judgments about the results of the assessment and the implications for programming?

7) What will be the locus of decisions for the evaluation of the assessment data? For example, will the evaluative judgments made be relevant to policies regarding students, faculty, programs, departments, or institutions?

8) What will be the uses of the evaluative judgments made based on the assessment data? Will the data and judgments be used for more summative purposes such as assessing worth of programs or formative purposes such as improving program functioning and quality?

Obviously, these eight parameters are not orthogonal dimensions and discussion of one parameter has implications for the discussion of other parameters. In the same fashion, it should be noted that policy decisions about one parameter may

constrain or influence decisions about other parameters (Alexander and Stark, 1986). This paper will focus on the purpose and use parameters, although the process-oriented view of assessment presented may have more implications for the technical parameters than the policy parameters.

A Process-Oriented View of Assessment

The purpose and use parameters are closely related and Alexander and Stark (1986) suggest that a consideration of the linkage between these two dimensions of assessment is crucial to any discussion of assessment. As noted above, there are many purposes and uses of assessment. This paper presents a relatively circumscribed view of the purpose and use of assessment in postsecondary settings. Specifically, the goal of this paper is to develop a process-oriented view of assessment that assumes that one of the most important purposes of assessment is to improve teaching and learning in postsecondary settings. Given this purpose, one of the best uses of assessment data is to focus on specific recommendations to improve the teaching and learning process. Accordingly, assessment in this view will have to concentrate on student, faculty, and classroom level data and the processes of learning and teaching. Although this type of data may not readily serve institutional, State, or Federal policy interests, it should serve the interests of faculty and students. As Scarr (1981) has stated, "testing should always be used in the interests of the children tested" (p. 1159). In addition, assessment should serve the interests of college faculty.

Obviously, college students and faculty are going to have a variety of interests and goals (e.g., research and publication on the part of the faculty and attainment of credentials for students). However, the process-oriented view of assessment assumes that most students would like to learn something in college and most faculty members would like to facilitate students' learning through teaching. How then can assessement serve these teaching and learning interests? The remainder of this paper proposes a process-oriented framework to address the use of assessment to improve teaching and learning in post-secondary settings.

Student Learning and Progress in Postsecondary Settings

Of course, an interest in teaching and learning begs the question, learning and teaching of what? As McKeachie, Pintrich, Lin and Smith (1986) note, many college faculty members have traditionally emphasized the goal of communicating knowledge to students. However, knowledge of facts and principles involves much more than memorization of isolated bits of information, concepts, and theories. Students need to have this material organized in a meaningful fashion so that they can use it in a variety of contexts. The ability to use information in different contexts is one aspect of a superordinate goal of life-long learning that most postsecondary institutions would subscribe to

either explicitly or implicitly. The goal of life-long learning implies that college students should learn not only knowledge of content, but also strategies for processing information and applying it critically to problems as well as the motivation to continue learning (McKeachie et al., 1986). In this paper, these four aspects of student learning, that is, knowledge, learning strategies, critical thinking and problem solving, and motivation, serve as the domains of interest for assessing student progress.

The notion of student progress implies development and change and, in an assessment context, measurement of that change. The measurement of change is a complicated technical issue that has a long history in educational research (e.g., Cook & Campbell, 1979; Cronbach and Furby, 1970; Kenny, 1979; Linn, 1986b; Rogosa, Brandt, & Zimowski, 1982) and will not be discussed here beyond noting that measurement of change generally implies multiple waves of measurement. Rather, the discussion will focus on a more theoretical analysis of student progress in terms of the psychology of the learner and the psychology of instruction.

There are a variety of developmental and psychological models that researchers have used to conceptualize student learning and development. Lerner (1986) has suggested that there are four very general theoretical approaches to psychological development, a stage theory approach, a differential approach, an ipsative approach and a behavioral approach. After a brief summary of these four approaches and their relevance to teaching, learning, and assessment in postsecondary settings, a fifth approach, an information processing view will be presented in detail.

The Stage Theory Approach. The stage theory approach to development is one of the most popular models and is often termed the developmental or classical approach (Lerner, 1986). Freud, Piaget, Erikson, and Kolhberg all proposed stage theories to explain human development. In the postsecondary area, Perry (1970) has formulated one of the most popular developmental models of college student thinking. Although stage models often provide useful ways of describing broad, general patterns in development, there are a number of theoretical and applied problems with them that reduce their utility as models for designing assessment tools of teaching and learning. First, it is not clear that development follows a fixed, universal, and invariant sequence of stages as implied by most stage models (cf. Siegler, 1986). Second, although most stage models assume that there are qualitative changes in the structure and organization of the individual's behavior over time, recent research in cognitive psychology suggests that change may be more quantitative in nature. For example, research on knowledge-based theories of learning (e.g., Bransford, Sherwood, Vye, & Rieser, 1986; Chi, 1978) suggest that students' "level" of development may depend more on the quality and depth of their knowledge in a specific content area than general cognitive structures.

Accordingly, a college student well-schooled in physics may be at a higher level in Perry's scheme in the domain of physics, but at a lower level in the domain of English literature. This issue of domain-specificity of knowledge presents serious problems for the application of general stage models to the development of global assessment tools to measure student progress. Besides these theoretical and applied problems with stage models, most of them have not clearly articulated a specific model of teaching that would help faculty design instruction to facilitate student progress through the stages.

The Differential Approach. As Lerner (1986) points out, the differential approach is basically an empirical, atheoretical model that can be used by researchers of any theoretical persuasion to address the basic question: "How, over the course of development, do groups of individuals become sorted into subgroups on the basis of status and behavioral attributes?" Status attributes are basically demographic characteristics of individual (e.g., age, gender, race, SES, etc.), while behavioral attributes are defined as psychological dimensions of individuals (e.g., personality characteristics such as independence-dependence, extroverison-introversion as well as cognitive characteristics such as intelligence). This approach has been the traditional approach to the measurement of intelligence (Sternberg, 1985) and has been used by a variety of researchers in higher education to assess student outcomes of college (e.g., Astin, 1977; Ewell, 1983,1984,1985; Pace, 1979). Although this approach has provided a sophisticated methodology for assessing student progress, the lack of a strong theoretical framework limits its utility for addressing specific problems of teaching and learning. However, just as Sternberg (1985) has done in the case of intelligence, the methodology of the differential approach can be combined with a strong, theoretical information processing framework to construct assessment tools that are relevant to teaching and learning issues. This combined approach will be described in more detail later. It is important to note that the development of assessment tools should be based on strong theoretical models of student learning, not just comprehensive listings and taxonomies of student outcomes. Without a good theoretical base that links assessment information to student learning and instruction, it is not clear what the assessment information means or how it can be used to improve teaching and learning.

The Ipsative Approach. In contrast to the differential approach which is more concerned with developing nomothetic laws of interindividual behavior, the ipsative approach is more concerned with the search for idiographic laws (Lerner, 1986). Accordingly, the ipsative approach focuses on intraindividual regularities and changes in behavior. Although the differential and ipsative approaches differ in the locus of their search for devlopmental change, the ipsative approach is basically atheoretical and can be used by researchers from a variety of theoretical perspectives. The ipsative approach attempts to characterize individuals in terms of the attributes that comprise

their intraindividual repertoire of behaviors, including personality characteristics, cognitive skills, attitudes, and beliefs. In addition, the ipsative approach attempts to uncover the intraindividual organization of these attributes over time (Lerner, 1986).

In the postsecondary area, a portfolio method of assessment may reflect an informal ipsative approach. In the portfolio approach, the student assembles a variety of products that represent their performance over the course of college. Generally, the individual student has control over the selection of products included in the portfolio. Examination of the contents of the portfolio can provide insight into the relative importance and organization of different attributes for the individual student. In contrast to the portfolio method, a more formal ipsative approach would involve the testing of the student using a variety of instruments, but instead of comparing the student to a set standard or other students as in the differential approach, the data could be analyzed using cluster analysis to construct individual profiles. These individual profiles could then be used to make inferences about the relative importance and organization of various attributes for the individual student. This focus on the individual student would provide useful information on intraindividual student progress in learning, but the lack of interindividual information results in the ipsative approach being less useful for the improvement of teaching.

The Behavioral Approach. Although older behavioral theories did not provide a role for cognition, newer models (e.g. Bandura, 1982) have postulated that the link between an environmental context and an individual's behavior is mediated by a variety of cognitive processes such as attention, self-regulation and monitoring, as well as motivation (Lerner, 1986). Accordingly, in newer behavioral approaches internal processes are not defined out of existence, but rather given central roles in the theory. However, most behavioral theories still prefer objective, behavioral measurement of these internal processes. In contrast to the qualitative change notions of stage theory approaches to development, developmental changes in these internal cognitive processes are assumed to be quantitative in nature. In addition, behavioral approaches are more concerned with nomothetic principles of development rather than idiographic concerns. In an assessment context, behavioral approaches often result in the specification of behavioral objectives and standards or criteria for students to meet on these objectives. As the history of educational testing suggests, this approach may provide useful information about the relative performance of students in a variety of contexts, but it does not necessarily provide useful information for teachers or students, or for those who would like to improve teaching and learning (e.g., Haney, 1984; Pintrich, Cross, Kozma, & McKeachie, 1986). In addition, as Bok (1986) points out, testing programs driven by behavioral objectives may be difficult to implement in many college settings because of the belief that no single set of goals can encompass

all the values of a liberal education, besides the practical issues concerning the enormous amount of time needed to develop common goals and the lack of agreement about what teaching methods to use to attain these goals.

All four of these approaches have something to offer us in considering how assessment can be used to improve teaching and learning. At the same time, they do not provide a strong enough, nor relevant and applicable, theoretical framework for understanding teaching and learning. Paralleling recent work in cognitive and instructional psychology (e.g., Pintrich et al., 1986), an information processing approach is adopted as a promising framework to discuss assessment, teaching and learning in postsecondary settings. This framework is described in more detail below and includes relevant aspects from the four previous approaches.

The Information-Processing Approach

As Pintrich et al., (1986) have pointed out, one advantage of adopting a general information processing approach is that it provides a common language for a wide range of educational phenomena. Not only are most experimental, social, developmental, instructional, and differential psychologists now using the same paradigm, but so are researchers on educational media and technology, instructional design, and classroom teaching. The confluence of all these different types of researchers who find information processing theory a useful theoretical framework for addressing psychological questions provides a rich store of knowledge to draw upon in discussing teaching and learning.

In contrast to stage models but similar to behavioral models, information processing views of development assume that student progress can best be characterized by quantitative not qualitative change in students' thinking or behavior. Although both differential and ipsative approaches have been used in information processing models, many information processing models have relied on a differential approach and attempted to sort people into different subgroups based on their knowledge base (e.g., the research on experts and novices, i.e., Chi, 1978) or on their use of strategies (e.g., the research on learning strategies and metacognition, i.e., Weinstein and Mayer, 1986). The information processing approach has been applied to student learning in a variety of ways. The following sections discuss four aspects of student learning that are relevant to teaching and learning in postsecondary settings. The four aspects are student knowledge, learning strategies, critical thinking and problem solving, and motivation.

Students' Knowledge. Many current information processing models of learning stress the importance of the nature and organization of the students' knowledge base for performance. For example, studies of experts' versus novices' memory for chess positions or dinosaur's names have demonstrated the importance of

knowledge for performance (Chi, 1978). In addition, studies of more school-related areas like reading, mathematics, physics, computer programming, and social science have all shown that effective problem solving in these domains depends on the students' knowledge in that domain (Bransford, Sherwood, Vye, & Rieser, 1986). These information processing views of student knowledge stress the organization of the students' knowledge into cognitive structures, schemata, cognitive representations, production systems, or propositional networks, to name a few of the labels used to describe the organization of student knowledge (McKeachie et al., 1986).

Although there are important theoretical differences among these labels, there are a few commonalities that are relevant for assessment issues. First, these knowledge structures are assumed to be internal, cognitive representations of the subject area. The students' structures are based on the content and structure of the information presented in class, but students actively create their own structures in the process of learning and integrating new course information with prior knowledge. These cognitive structures help students store and organize new information, thereby guiding future perceptions and learning. Finally, although these cognitive structures are internal, cognitive representations, they are can be inferred and described by objective measures (McKeachie et al., 1986) and represent the "students' public understanding of a discipline" (Shavelson, 1983).

How is this information processing, cognitive structure view of student knowledge relevant to college teaching and learning? First, it is likely that college students learn a great deal of content knowledge in different courses and that students' level of content knowledge is assessed rather well by the instructors' tests and final examinations. However, the relevance of the information processing approach lies in the assumption that it is not just the amount or quantity of knowledge that a student possesses, but the qualitative organization and structure of that knowledge that is important for future performance. As McKeachie et al. (1986) point out, if an important goal of higher education is to foster life-long learning, then the student must be able to use the knowledge gained in a course after the final examination. This is not likely to happen if the information is just a compilation or accumulation of unrelated facts, principles, and theories. The information must be organized and elaborated in a meaningful fashion. The difficult and important issues of how students acquire meaningful knowledge and how instruction can facilitate that process are beyond the scope of this paper, but the assessment of student knowledge and how that assessment information may be used to improve teaching and learning is central to the discussion.

There are a variety of methods that have been used in attempts to measure students' cognitive or knowledge structures. McKeachie et al. (1986) provide a brief overview of some of the

most common methods. There are two general approaches, a direct one and an indirect one. In the variations of the direct method, students are asked to arrange key concepts or propositions from the course in some structured, spatial manner to reflect the students' understanding of the course material. This can include drawing maps, networks, or other pictorial, graphic representations of the relationships among the concepts. The resulting product is scored according to preset criteria and used as a measure of students' cognitive structure. In contrast, the indirect methods do not have the students produce an actual spatial representation of the relationships among the concepts. Indirect methods include word association, sorting, ordered-tree, and interview techniques. Basically, these indirect methods infer students' cognitive structures from how the students sort or arrange the lists of concepts. Different scaling methods are then used to assign scores to students' arrangements of the concepts.

These techniques can be used to examine the development of students' cognitive structures over the course of a term. For example, in a series of studies at Michigan, my colleagues, Bill McKeachie, Yi-Guang Lin, and Moshe Naveh-Benjamin, have examined the relationship between students' cognitive structures and achievement in the class as well as students' progress toward more organized and hierarchical structures over the semester (e.g., Naveh-Benjamin, McKeachie, Lin, & Tucker, 1986). They found that students' cognitive structures were correlated with performance in the class, but not to such an extent that they seemed to be measuring the same construct. In addition, students' structures became more organized, more elaborated, and more similar to the instructor's cognitive structure over the course of the semester.

This type of measure has a number of applications for assessment and teaching and learning. First, it can be used by instructors to assess students' cognitive structures for the course material. It provides a measure of student performance in the course that is related to achievement on examinations, but reflects students' organization of the material, not just their recall of the material. This information can be used by instructors to assess their efficacy in teaching the organization of the course content. In addition, we have used the method as a teaching tool in the classroom to improve student learning. We have found that the construction of cognitive structures by students and the explicit comparison of student structures with faculty structures facilitates discussion of course content. For example, the discussion of different models for understanding and organizing the course material seems to help students comprehend some of the underlying relationships among the course concepts. In addition, it assists students in seeing that there are different ways of accurately representing course material (e.g., there is no one right answer). Finally, this method can help faculty members improve their teaching by making explicit their goals and preferred structures for the organization of course material. For example, in our discussions of cognitive structure

with faculty members from other disciplines (e.g., English composition and literature, biology, ecology, sociology), the method has provided a means for clarifying the goals of their course and their instructional strategies for achieving these goals. In several cases, these discussions have highlighted for the faculty the lack of accordance between their goals for the course and their instructional strategies for achieving them. One outcome of these informal discussions is that faculty members can then decide if they want to adjust their instructional goals and strategies to make them more consonant.

Although the content and structure of students' knowledge obviously plays an important role in their learning, it may not be sufficient for all learning situations and effective problem solving (Pintrich, et al., 1986). Educators at all levels have become increasingly concerned about generalizable cognitive skills such as those for processing information as well as critical thinking and problem solving. The next two sections focus on these generalizable cognitive skills.

Students' Learning Strategies. As Weinstein and Mayer (1986) point out, recent research on teaching and learning has focused on the active role of the learner in student achievement. Obviously, the subject matter content students know when taking on a new task will influence their performance. Accordingly, theories about prior knowledge and cognitive structures are important components of a theory of learning. Many of these knowledge-driven models, however, do not address how the students orginally acquired that knowledge. While there are a variety of theories about the acquisition of knowledge that are beyond the scope of this paper, the area of research on student learning strategies that deals with how students acquire and modify their knowledge base will be discussed.

Weinstein and Mayer (1986) in their review of the learning strategies literature suggest that learning strategies influence the way students process information as they acquire new knowledge and modify old or prior knowledge. They suggest that there are four components of information processing that are affected by learning strategies. The four processes are: selection, acquisition, construction, and integration. There are other models of this process (e.g., Corno and Mandinach, 1983; Sternberg, 1985) that use somewhat different terms, but the general concern is how students acquire and integrate new information with old information.

The selection process involves how attention is controlled to select certain stimuli or information in the environment and transfer that information to working memory. The acquisition phase concerns the transfer of information from working memory to long term memory for permanent storage. This involves how the information is encoded by the student. In the construction phase the student actively builds connections between ideas and concepts in working memory. This construction involves the building of schemas and other organizational

frameworks for the information as discussed in the previous section. The last phase, integration, involves connecting this new knowledge with prior knowledge so that retrieval is facilitated in the future (McKeachie et al., 1986). This brief description of these four processes does not do justice to the complicated problem of how students process information, but it provides a context for the discussion of learning strategies.

There are many definitions of learning strategies. Weinstein and Mayer (1986) define learning strategies as any thought or behavior that a learner engages in while learning. This includes basic memory processes as well as general problem solving strategies. This is a very broad definition of learning strategies and encompasses almost all the researched cognitive processes. In contrast, Tobias (1982) has distinguished between microlevel strategies and macrolevel strategies. For Tobias, the microlevel strategies concern the basic cognitive processes such as attention and encoding of all information, while the macrolevel strategies such as reviewing, notetaking, and comprehension monitoring concern the processing of instructional input. This distinction parallels Sternberg's (1985) distinction between performance components and metacomponents.

In terms of teaching and learning, the more molar level of analysis provided by a focus on macrolevel strategies is more relevant (McKeachie et al., 1986). This choice is made on theoretical, methodological, and practical grounds. First, a number of researchers (e.g., Paris, Lipson, & Wixson, 1983) have limited their definition of learning strategies to cognitive processes that are intentional and under the control of the learner. Some of the basic attentional and memory processes are not really under the control of the learner; they are part of every individual's basic information processing equipment and are elicited automatically by various tasks (see Sternberg, 1985).

Second, the basic cognitive microprocesses are difficult to measure unless experimental designs with carefully specified tasks are combined with the collection of reaction time data. This is clearly not a practical option for most people concerned with assessment of teaching and learning in college settings. In addition, it is not clear that some of the experimental tasks used by cognitive psychologists in the laboratory have much ecological validity when applied to the classroom setting. Accordingly, in attempting to assess student learning strategies we will have to rely on other methods (e.g., self-report measures) that may decrease construct validity but increase external validity (Cook & Campbell, 1979).

In attempting to assess students' learning strategies, there are a variety of taxonomies available for describing and classifying learning strategies (cf., Dansereau, 1985; Pressley, 1986; Weinstein & Mayer, 1986). My colleagues and I (McKeachie et al., 1986) have recently grouped them into the three general categories of cognitive, metacognitive, and resource management strategies. Table 1 displays these three categories and lists

examples of the types of strategies included in each category. Basically, the cognitive strategies are assumed to help students encode new material and facilitate the organization and retreival of information. The metacognitive strategies assist students in planning, regulating, monitoring, and modifying their cognitive processes. The resource management strategies help students control the various resources they have available to them (i.e., time, effort, outside help) for performing various academic tasks.

The problem of how to assess these learning strategies is not a new issue. There have been a number of attempts to develop measures of study skills and learning strategies (cf., Brown, 1964; Brown & Holtzman, 1967; Carter, 1958; Christensen, 1968; Goldman & Warren, 1973). However, as Weinstein and Underwood (1985) have pointed out, there are several problems with these instruments. Most importantly, many of them have no underlying theoretical framework. Items are included concerning traditional areas of study skills such as notetaking, time management, work habits, and attitudes, but there are few items on how students actually learn or process material. More recent approaches like Weinstein's (forthcoming) Learning Activities and Study Strategies Inventory (LASSI) are based on current instructional and cognitive research as well as information processing approaches to learning. In our own work at Michigan, we have been developing a similar self-report questionnaire (the Motivated Strategies for Learning Questionnaire or MSLQ) based on an integration of motivational and information processing theories of learning. The higher order cognitions that are measured by these learning strategies instruments (e.g., metacognition, active learning, self-regulated learning) should be intended or, at least, unintended outcomes of a college education. As such, it is important to incorporate constructs based on these information processing views of student learning in postsecondary assessment programs.

How can instruments developed to measure students' learning strategies help improve teaching and learning? First, although it is assumed that students' learning strategies will improve over the course of their college career, it is not clear that most college courses will have a direct influence on students' learning strategies. (Of course, this excludes various study skills or learning to learn college courses that are designed to explicitly teach learning strategies.) Accordingly, in contrast to the cognitive structure measurement data, the usual aggregate data generated from general learning strategy questionnaires may not be readily usable by individual faculty members. The aggregate data could be utilized at the institutional level to ascertain general skill levels of students and then be used to guide curriculum offerings. For example, if a community college finds that many of their students have generally low levels of learning strategies, they could decide to offer specific courses to improve learning strategies. This use of the differential approach to learning strategies data could help students improve their learning skills so they could benefit

more from college. At an individual student level, we have found that providing students with individual feedback about the level of their learning strategies motivates students to think about their skills in a way they haven't in the past (e.g., that these strategies are skills that are learnable, and that academic performance is not totally determined by general ability or intelligence). In addition, since the feedback is based on classroom norms not national norms, many students find the information usable and are able to act on it (e.g., change their study behaviors or resource management behaviors).

In contrast, to these differential approaches, the data also could be used in an ipsative format that would be more useful to individual faculty members. For example, an instructor could track how different types of students (classified on the basis of the learning strategy data) perform in the class over the course of the semester. At the end of the semester, the instructor could compare the learning strategy patterns of students who did well in the course and those that did not. This use of ipsative data would provide more information to the instructor about the nature of student performance in the class. In addition, patterns in the data might reveal attribute-treatment interactions between types of students and different methods of instruction (e.g., Corno & Snow, 1986). For example, in a biology class, one type of student may do very well on examinations on lecture material, while another type of student does very well on the laboratory exercises. Although the practical and logistical aspects of using ipsative data for improving an individual course are many, ipsative data on student learning strategies could be useful to faculty members by helping them diagnose the locus of learning problems in their classroom. That is, the data could help faculty members assess whether problems in student learning in their class are due to student characteristics such as low levels of motivation or learning skills or characteristics of the class including the instructional methods and tasks.

Students' Critical Thinking and Problem Solving. Critical thinking and problem solving is currently a "hot" topic in education from elementary through secondary levels up into postsecondary education. There are a number of books appearing on the topic (e.g., Baron & Sternberg, 1987; Chipman, Segal, & Glaser; 1985; Nickerson, Perkins, & Smith, 1985; Segal, Chipman, & Glaser, 1985) as well as articles (e.g., the entire winter 1984 issue of Review of Educational Research deals with problems in the teaching and learning of reasoning skills). Although this attention to the topic is interesting and exciting, there is a great deal of theoretical confusion concerning the nature of critical thinking and problem solving. Researchers and practitioners are addressing the topic from a multitude of perspectives and the constructs are not clearly defined. It is beyond the scope of this paper to attempt to clarify the theoretical and definitional issues, but several other general issues will be addressed that are relevant to assessment of critical thinking in postsecondary education.

In an excellent overview of the critical issues involved in attempting to teach general cognitive skills, Glaser (1984) suggests that there are five issues that must be considered in future research:

1) Can general cognitive skills be taught?

2) How can current knowledge-based models of cognition be applied to the teaching of general problem solving skills?

3) How can instruction best be designed to foster general cognitive skills?

4) How can cognitive skills learned in one domain, be transferred to another domain?

5) How can we assess the effectiveness of our attempts to teach critical thinking or other general cognitive skills?

There may be a variety of answers to these questions, depending on the theoretical orientation of the individuals involved (e.g., faculty, institutional researchers, administrators, etc.), but consideration of these issues is important for building assessment programs of critical thinking. By having to think about these issues, the individuals have to consider their definitions of critical thinking and problem solving as well as their implicit theories about how students learn these skills and how teachers might teach these skills. Ideally, these deliberations would lead to the specification of a local model or theory of how instruction will lead to critical thinking and problem solving. Regardless of the level of the local model (e.g., course, departmental, institutional), the development of a local model of critical thinking would help to delineate how the "independent" variables of course tasks and activities, curriculum offerings, or institutional dimensions theoretically influence the "dependent" variable of students' critical thinking. Assessment programs could then be designed to test this local model. The data from this kind of assessment program would provide very useful feedback to instructors (or other program designers) about the relative efficacy of different aspects of their program. This information could then be used to improve or redesign the critical thinking instructional program or curriculum.

This suggestion for the development of local models of instruction and thinking parallels recent work by methodologists working in the area of program evaluation. For example, Corday (1986) suggests that traditional conceptualizations of treatments and quasi-experiemental designs are no longer useful in program evaluation research. Corday (1986) suggests that evaluators should build conceptual models that specify student entry characteristics (exogeneous variables), treatment variables (including both theoretical aspects as well as practical aspects such as treatment fidelity and implementation), mediating

variables, and finally outcome variables. The specification of these more theoretical models for postsecondary assessment programs will help us avoid simplistic input-output models of college impact that, at worst, may be misleading and, at best, do not provide useful information that can be readily used to improve teaching and learning.

For example, a focus on treatment implementation and mediating variables in postsecondary assessment programs of critical thinking would highlight the importance of considering instructional and cognitive process variables as mediators of the impact of college on critical thinking. This focus would help us avoid problems of assessing the effects of programs that can not be realistically expected to result in student progress on a global measure of critical thinking (e.g., finding that a specific college curriculum has no influence on students' score on the Watson-Glaser Critical Thinking Test). As the research on mainstreaming (e.g., Leinhardt & Palley, 1982) and open education (Marshall, 1981) has shown us, it is not the label of the placement or placement per se that influences student learning in educational settings, but what happens to the student in those settings. In the same manner, at the postsecondary level it may not matter how the curriculum is labeled (e.g., interdisciplinary curriculum for critical thinking) or what the course is titled (e.g., critical thinking), but what happens to the students in those classes and what they are asked to do that will influence their learning (Doyle, 1983; Pintrich et al., 1986). This process-oriented view may not serve other policy interests, but it should result in increased knowledge about how students learn to think critically and how we can improve our instructional efforts in this area.

Students' motivation. One process or mediating variable that is almost always left out, ignored, or poorly conceptualized in most cognitive models of students' learning, critical thinking, or problem solving is motivation. It is clear, however, that initial acquisition of knowledge as well as transfer of general cognitive skills across different content domains requires a motivated learner (McKeachie et al., 1986). One problem with past models of motivation is that they have not been readily applicable to cognitive models of learning. For example, motivational models based on psychodynamic, simple drive, or humanistic theories have not been useful in conceptualizing student learning. In addition, motivational models based on basic attitude constructs (e.g., values, liking, or interest) have proven too simplistic. In recent years, however, cognitive reformulations of achievement motivation theory have revitalized motivational research and suggested productive relationships between research programs in motivation and cognition (Pintrich et al., 1986). Figure 1 displays one version of a general expectancy-value model of motivation that can be used to conceptualize the interactions between motivational, cognitive, and instructional variables. A brief overview of Figure 1 is provided here, for more detail see McKeachie et al. (1986).

There are two general paths in the model, an expectancy path and a value path. The expectancy path is along the bottom of Figure 1 and includes expectancies, perceived competence, test anxiety, perceptions of task difficulty, and students' beliefs about efficacy, control, and outcome. This aspect of the model subsumes general motivational constructs like self-concept and self-efficacy. The basic outcome of this path in the model is the student's belief or expectation for success. The model assumes that students with generally high expectations for success for a specific task (e.g., an exam, an assigned paper, a course) will be more involved in the task and persist longer on the task in the face of difficulty. In contrast, students with low expectations for success should be less involved and give up more easily. In addition, the model predicts a relationship between expectancy for success and cognition by suggesting that high expectation students will be more likely to recruit more effective cognitive strategies to help solve the task than students with low expectations.

At the same time, expectancy for success is not the only important motivational component in determining student involvement with the task. The task value path, along the top of Figure 1, includes task value and student goal components. The task value component includes three aspects, atttainment value, interest value, and utility value. Attainment value refers to the students' perception of the task's ability to provide a challenge or to confirm a salient aspect of the self such as competence. For example, a student who believes she is smart and perceives doing well in a chemistry course as a challenge would have a high attainment value for the course. Interest value includes the students' intrinsic interest in the content of the task (e.g., the student enjoys chemistry). Many college courses have implicit goals of increasing students' intrinsic interest in the discipline. In contrast to the "means" or process aspect of interest value, utility value concerns the "ends" or instrumental aspect of a task, that is, the student's perception of the utility of the task for facilitating a specific goal. For example, a student may have no inherent interest in chemistry, but because she has a goal of becoming a doctor, and chemistry is a required course for premeds, the course has a high utility value for her. This recognition that tasks may have utility value as well as intrinsic value for students lessens the need to assume that all instruction has to increase students' intrinsic interest in learning or education.

The other aspect of the task value path is the student's goals. These goals can include both long-term (e.g., career, life) and short-term goals (e.g., goals for a specific course, paper, or exam). These goals help determine the student's perception of the value of a task and their overall choice of tasks (e.g., selecting certain courses). The value and expectancy components interact to produce student involvement in this model. Both components are assumed to be important predictors of student involvement in college. Assesment models that only include expectancy components like self-concept, self-

efficacy, or expectancy for success or models that only include value components like interest or utility value will not be as powerful as models that include both expectancy and value components. It is important that assessment programs incorporate these motivational and value components in their assessment systems. If assessment systems only focus on narrowly defined cognitive aspects of student performance, they risk missing some of the more "intangible" aspects of the college experience in terms of students' intrinsic interest and value for learning (cf., Bennett, 1986).

The assessment of motivational constructs has been a continuing problem in motivational research. Older models used projective techniques and behavioral observation, but newer cognitive models have relied on self-report instruments (McKeachie et al., 1986). Although there are problems with self-report instruments (e.g., Nisbett & Wilson, 1977), they can be used to tap students' expectancies and values in a reliable fashion. As discussed previously in the learning strategies section, data from this type of assessment could be used by faculty members to determine the relative effects of their course on students' motivational patterns. For example, by tying the questionnaire items to specific course content, the instructor could determine the development or progress in students' interest value for the course material or their perceived competence. In addition, using the motivational data ipsatively, faculty members could look at the interactions between students' motivational patterns and their performance in the course. As noted before, this type of information would be very helpful to faculty members in redesigning or improving their course.

Summary and Policy Implications

In summary, I would like to make five general points.

1) As Haney (1984), Linn (1986), and Sternberg (1985) have all noted, new assessment or testing programs must be more closely tied to instructional concerns as well as current theory on cognitive models of student learning. Accordingly, if researchers and policy makers are serious about the goal of improving postsecondary teaching and learning and plan to use assessment programs as one tool to reach this goal, then the assessment program must be linked to a strong, theoretical framework of student learning, motivation, and instruction. Accordingly, assessment programs that are basically atheoretical in nature, that include a variety of items tapping a diversity of student outcomes, and are not linked to concerns of teaching and learning, will not be very useful in improving teaching and learning in higher education. Although assessment programs that are atheoretical and based on general "dustbowl empiricism" methods may have some administrative utility, they generally do not provide information that is readily usable for instructional improvement efforts.

2) Although I believe the general information processing models presented in this paper are the most relevant, there may be other models that are useful in designing assessment programs. However, it is important to note that these other models should be concerned not only with the psychology of student learning, but also with the psychology of instruction.

3) As Corday (1986) has pointed out, simple input-output models are not adequate for most program evaluation tasks. Assessment programs need to focus not only on student outcomes, but also instructional processes and other mediating constructs.

4) Given the goal of improving teaching and learning, the process-oriented model presented in this paper implies that many of the policy parameters and decisions discussed by Alexander and Stark (1986) will be focused at the local institutional or faculty level. Accordingly, colleges need to develop the internal capabilities to implement the types of assessment programs suggested here. For example, institutional researchers and faculty need to be trained in the development and use of these programs.

5) Finally, given this local policy focus, State and Federal policy makers may only have an indirect role to play in the development of the suggested assessment programs. As Bennett (1986) has pointed out, "The federal government cannot and should not play the primary role in the assessment of higher education." (P. iii.) Accordingly, the Federal Government should foster assessment activities that help to improve higher education by focusing the discussion on how assessment programs can be used to improve teaching and learning.

References

Adelman, C. (1986). To imagine an adverb: Concluding notes to adversaries and enthusiasts. In C. Adelman, (Ed.), Assessment in American Higher Education: Issues and Contexts (pp. 73-82). Washington, D.C.: U.S. Government Printing Office.

Alexander, J. & Stark, J. (1986). Focusing on student academic outcomes: A working paper. Paper prepared for NCRIPTAL, The University of Michigan, Ann Arbor, Michigan.

Astin, A. (1977). Four Critical Years: Effects of Colleges on Beliefs, Attitudes, and Knowledge. San Francisco: Jossey-Bass.

Bandura, A. (1982). Self-efficacy mechanism in human agency. American Psychologist. 37, 122-147.

Baron, J., & Sternberg, R. (1987). Teaching Thinking Skills. New York: Freeman.

Bennett, W.J. (1986). Foreword. In C. Adelman (Ed.), Assessment in American Higher Education: Issues and Contexts (pp. i-iii). Washington, D.C.: U.S. Government Printing Office.

Bok, D. (1986). Higher Learning. Cambridge, MA: Harvard University Press.

Boyer, E. (forthcoming). College: The Undergraduate Experience in America. New York: Harper and Row.

Bransford, J., Sherwood, R., Vye, N., Rieser, J. (1986). Teaching thinking and problem-solving: Research foundations. American Psychologist, **41**, 1078-1089.

Brown, W. (1964). Effective Study Test. San Marcos, Texas: Effective Study Materials.

Brown, W., & Holtzman, W. (1967). Survey of Study Habits and Attitudes. New York: The Psychological Corporation.

Carter, J. (1958). California Study Methods Survey. Monterey, CA: California Test Bureau.

Chi, M. (1978). Knowledge structures and memory development. In R. Siegler (Ed.), Children's Thinking: What Develops? (pp.73-96). Hillsdale, N.J.: Erlbaum.

Chipman, S., Segal, J., & Glaser, R. (1985). Thinking and Learning skills. Research Issues (Vol. 2). Hillsdale, N.J.: Erlbaum.

Christensen, F. (1968). College Adjustment and Study Skills Inventory. Berea, Ohio: Personal Growth Press.

Cook, T. & Campbell, D. (1979). Quasi-Experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNally.

Corday, D. (1986). Quasi-experimental analysis: A mixture of methods and judgment. In W. Trochim (Ed.), Advances in Quasi-Experimental Design and Analysis: New Directions in Program Evaluation Vol. 31 (pp. 9-28). San Francisco: Jossey-Bass.

Corno, L., & Mandinach, E. (1983). The role of cognitive engagement in classroom learning and motivation. Educational Psychologist, 18, 88-100.

Corno, L., & Snow, R. (1986). Adapting teaching to individual differences among learners. In M. Wittrock (Ed.), Handbook of Research on Teaching (pp. 605-629). New York: Macmillan.

Cronbach, L., & Furby, L. (1970). How we should measure "change"--or should we? Psychological Bulletin, 74, 68-80.

Dansereau, D. (1985). Learning strategy research. In J. Segal, S. Chipman, & R. Glaser (Eds.), Thinking and Learning Skills: Relating Instruction to Research (Vol. 1), (pp. 209-240). Hillsdale, N.J.: Erlbaum.

Doyle, W. (1983). Academic work. Review of Educational Research, 53, 159-199.

Ewell, P. (1983). Information on Student Outcomes: How to Get It and How to Use It. Boulder, CO: National Center for Higher Education Management Systems.

Ewell, P. (1984). The Self-Regarding Institution: Information for Excellence. Boulder, CO: National Center for Higher Education Management Systems.

Ewell, P. (1985). (Ed.), Assessing Outcomes: New Directions for Institutional Research, No. 47. San Francisco: Jossey-Bass.

Glaser, R. (1984). Education and thinking: The role of knowledge. American Psychologist, 39, 93-104.

Goldman, R., & Warren, R. (1973). Discriminant analysis of study strategies connected with college grade success in different major fields. Journal of Educational Measurement, 10, 39-47.

Haney, W. (1984). Testing reasoning and reasoning about testing. Review of Educational Research, 54, 597-654.

Hartle, T. (1986). The growing interest in measuring the educational achievement of college students. In C. Adelman (Ed.), Assessment in American Higher Education: Issues and Contexts (pp.1-12). Washington, D.C.: U.S. Government Printing Office.

Kenny, D. (1979). Correlation and Causality. New York: John Wiley.

Leinhardt, G., & Pallay, A. (1982). Restrictive educational settings: Exile or haven. Review of Educational Research, 52, 557-578.

Lerner, R. (1986). Concepts and Theories of Human Development. New York: Random House.

Linn, R. (1986a). Educational testing and assessment: Research needs and policy issues. American Psychologist, 41, 1153-1160.

Linn, R. (1986b). Quantitative methods in research on teaching. In M. Wittrock, (Ed.), Handbook of Research on Teaching (pp. 92-118). New York: Macmillan.

Loacker, G., Cromwell, L., & O'Brien, K. (1986). Assessment in higher education: To serve the learner. In C. Adelman (Ed.), Assessment in Higher Education: Issues and Contexts (pp. 47-62). Washington, D.C.: U.S. Government Printing Office.

Marshall, H. (1981). Open education: Has the term outlived its usefulness? Review of Educational Research, 51, 181-192.

McKeachie, W., Pintrich, P.R., Lin, Y., & Smith, D. (1986). Teaching and learning in the college classroom: A review of the research literature. Paper prepared for NCRIPTAL, The University of Michigan, Ann Arbor, Michigan.

Mortimer, K., Astin, A., Blake, J., Bowen, H., Gamson, Z., Hodgkinson, H., & Lee, B. (1984). Involvement in Learning: Realizing the Potential of American Higher Education. Washington, D.C.: U.S. Government Printing Office.

National Comission on Excellence in Education (1983). A Nation at Risk: The Imperative for Educational Reform. Washington, D.C.: U.S. Government Printing Office.

Naveh-Benjamin, M., McKeachie, W.J., Lin, Y., & Tucker, D. Inferring students' cognitive structure and their development using the "ordered-tree technique". Journal of Educational Psychology, 78, 130-140.

Nickerson, R., Perkins, D., & Smith, E. (1985). The Teaching of Thinking. Hillsdale, N.J.: Erlbaum.

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.

Pace, C. (1979). Measuring Outcomes of College. San Francisco: Jossey-Bass.

Paris, S., Lipson, M., & Wixson, K. (1983). Becoming a strategic reader. Contemporary Educational Psychology, 8, 293-316.

Perry, W. (1970). Forms of Intellectual and Ethical Development in the College Years: A Scheme. New York: Holt, Rinehart, and Winston.

Pintrich, P.R., Cross, D., Kozma, R., & McKeachie, W.J. (1986). Instructional psychology, Annual Review of Psychology, 37, 611-651.

Pressley, M. (1986). The relevance of the good strategy used model to the teaching of mathematics. Educational Psychologist, 21, 139-161.

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. Psychological Bulletin, 92, 726-748.

Scarr, S. (1981). Testing for children: Assessment and the many determinants of intellectual competence. American Psychologist, 36, 1159-1166.

Segal, J., Chipman, S., & Glaser, R. (1985). Thinking and Learning Skills: Relating Instruction to Research (Vol. 1). Hillsdale, N.J.: Erlbaum.

Shavelson, R. (1983). On quagmires, philosophical and otherwise: A reply to Phillips. Educational Psychologist, 18, 81-87.

Siegler, R. (1986). Children's Thinking. Englewood Cliffs, N.J.: Prentice-Hall.

Sternberg, R. (1985). Beyond IQ: A Triarchic Theory of Human Intelligence. Cambridge: Cambridge University Press.

Tobias, S. (1982). When do instructional methods make a difference? Educational Researcher, 11, 4-10.

Weinstein, C. (forthcoming). The Learning Activities and Study Skills Inventory (LASSI). The University of Texas, Austin, TX: H & H Publishing.

Weinstein, C., & Mayer, R. (1986). The teaching of learning strategies. In M. Wittrock, (Ed.), Handbook of Research on Teaching (pp. 315-327). New York: Macmillan.

Weinstein, C., & Underwood, V. (1985). Learning strategies: The how of learning. In J. Segal, S. Chipman, & R. Glaser, (Eds.), Thinking and Learning Skills: Relating Instruction to Research (Vol. 1), (pp. 241-258). Hillsdale, N.J.: Erlbaum.

Table 1

A TAXONOMY OF LEARNING STRATEGIES

I. COGNITIVE STRATEGIES

Basic Tasks (e.g., memory for lists)	Complex Tasks (e.g., text learning)
A. Rehearsal Strategies	
reciting list	shadowing copy material verbatim notetaking underlining text
B. Elaboration Strategies	
keyword method imagery method of loci	paraphrasing summarizing creating analogies generative notetaking question answering
C. Organizational Strategies	
clustering mnemonics	selecting main idea outlining networking diagramming

Table 1 (continued)

II. METACOGNITIVE STRATEGIES

A. Planning Strategies
- setting goals
- skimming
- generating questions

B. Monitoring Strategies
- self-testing
- attention-focus
- test-taking strategies

C. Regulating Strategies
- adjusting reading rate
- re-reading
- reviewing
- test-taking strategies

III. RESOURCE MANAGEMENT STRATEGIES

A. Time management
- scheduling
- goal setting

B. Study environment management
- defined area
- quiet area
- organized area

C. Effort management
- attributions to effort
- mood
- self-talk
- persistence
- self-reinforcement

D. Support of others
- seeking help from teacher
- seeking help from peers
- peer/group learning
- tutoring

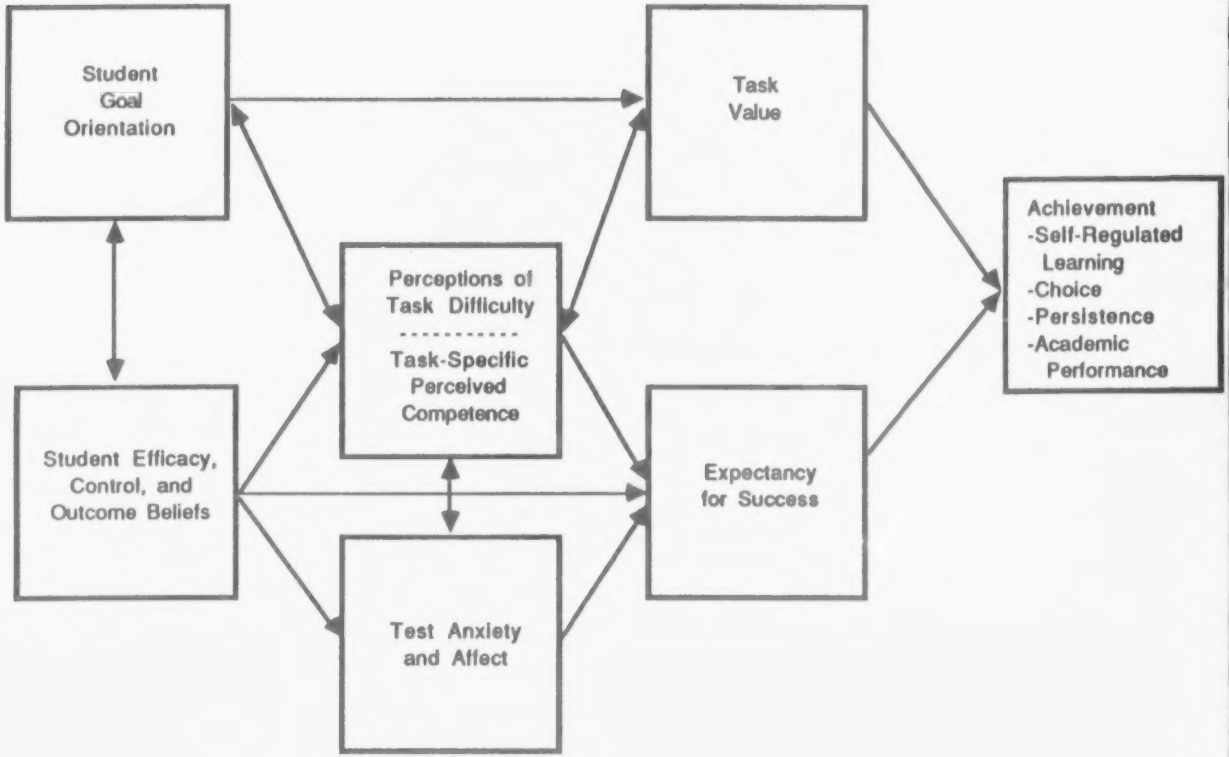


Figure 1. Components of Motivation

Aubrey Forrest
American College Testing Program

The increasing attention being paid to the assessment of outcomes in America's colleges and universities is a response to growing concerns about the quality of undergraduate degree programs. Specifically, there is an expanding need for postsecondary institutions to gather and provide pertinent information about student levels of skill proficiency, knowledge acquisition and value development. Such information can be used by the institutions to help students learn more and enhance institutional credibility.

While it is true that other "college outcomes" (e.g., the outcomes of community service and research activities) could become the foci of attention, current national concerns clearly center upon student learning outcomes.

It is also true that assessments of student learning could serve purposes other than the improvement of undergraduate education. Graduate and professional education could come under scrutiny or such assessments could be used in making decisions about individual students for purposes of admission, placement, progression and/or credentialing. The current new wave of concern is, however, aimed at the use of assessment for purposes of undergraduate program evaluation and improvement.

Various organizations and individuals have publicly recommended that undergraduate education be strengthened by evaluating the degree to which expected results (intended learning outcomes) are achieved by students and, by inference, by the colleges and universities. For example, while acknowledging that postsecondary institutions are evaluated by accrediting and licensing agencies, largely external groups, the Carnegie Foundation for the Advancement of Teaching suggested that the institutions should also employ evaluation criteria of their own to determine whether they are meeting educational goals (CFAT, 1977). At a 1978 conference sponsored by the Council on Postsecondary Accreditation, participants agreed that, when evaluating educational quality, accrediting bodies should focus on "educational outcomes"--the product of education rather than the process and the setting (COPA, 1979). In furtherance of this thought, the Association of American Colleges noted that the present system based on course grades and overall grade point average tells little about what students have actually accomplished (AAC, 1982). The National Commission on Excellence in Education stated that many readers of college transcripts remark, "We know what they take, but we don't know what they have learned." (NCEE, 1983.) The American Council on Education has urged that steps be taken to assure that college graduates are at least minimally proficient in the attributes of the broadly-educated person (Miller, 1978). The American College Testing

Program (ACT), through its College Outcomes Measures Program (COMP), has demonstrated the value of information on student learning outcomes in strengthening general education programs (Forrest, 1982). Jaschik has observed that "at the behest of legislature and governing boards, a growing number of public colleges and universities are using 'value added' tests and surveys in an attempt to measure what students learn in college." (Jaschik, 1985.)

A National Institute of Education Study Group suggested that it is "futile to engage in changing the curriculum and/or instructional approaches without knowing whether these changes have been successful." It stated that "a comprehensive assessment program will help faculty determine what works and what does not." It concluded that regular and periodic assessment is essential to effective learning--to identifying the strengths and weaknesses of an academic program (NIE, 1984).

The Association of American Colleges issued a report suggesting that there must be some hard evidence of student progress toward the goal of effectively functioning human beings, if a faculty is to know how well it is doing its job (AAC, 1984). The Southern Regional Education Board urged the use of comprehensive examinations or other kinds of evaluations (apart from course grades) to encourage faculty and students to develop the fundamental outcomes of reading, writing, mathematics and critical thinking (SREB, 1985). Education Secretary William J. Bennett sees assessment of student learning by colleges as the "surest way to turn the lofty statements of college catalogs into actual classroom practice." (USED, 1986.) Under the sponsorship of the Education Commission of the States, a 21-person panel chaired by New Jersey Governor Thomas H. Kean observed that assessment should play an integral role in attempts to improve teaching and learning and institutional performance. It concluded that assessment can guide the development of students and institutions (ECS, 1986).

In a survey of 365 colleges and universities, 75 percent of the presidents (or other key officers) stated that they favor the development of student assessment procedures, 88 percent felt that the results of student assessment should be linked to analyses of institutional effectiveness, and 91 percent said that student assessment should be closely tied to efforts to improve instruction (ACE, 1986).

A recent study by the National Governors' Association (in which every state governor participated) points out that quality can be best assured when outcomes information about students, programs and institutions is regularly collected and used to improve undergraduate teaching and learning. Further, in order to motivate institutions to improve, assessment information should be made public and should be used as a basis for providing incentive funding to reward institutions that can demonstrate that they are effective in educating students (NGA, 1986). In a Carnegie Foundation for the Advancement of Teaching report,

Ernest L. Boyer has warned that academic leaders cannot ignore the pressure from political leaders to measure the results of college attendance (Boyer, 1986).

Running throughout these reports are assumptions that two types of assessment data are needed to assist in the improvement of undergraduate education: (1) measures of existing level proficiency for graduates and (2) measures of intellectual growth during college attendance. Information about abilities of graduates at the time of exiting college requires measurement at only that one time. To gain information about intellectual growth (so called "value-added assessment") requires testing of entering freshmen as well as those students near graduation. To accomplish this, many advocates of the value-added approach appear to be favoring longitudinal studies of the same students as they progress through college, though the cross-sectional design has its adherents.

Nearly all of the recent reports recommending the assessment of student learning outcomes strongly suggest that multiple techniques be utilized. For a variety of reasons, using only a single approach is seen as inadequate. Present instruments are not sufficiently valid and reliable, the span of abilities to be measured is highly varied, higher education is dealing with a highly diverse student population, educational goals and objectives vary from college to college, etc.

For many of these types of reasons, the NIE Study Group recommended a comprehensive assessment program, including standardized tests, essays, interviews, portfolios and performance examinations (NIE, 1984). In its survey, the American Council on Education found that 83 percent of the presidents (or other key officers) believe that the following measures would be appropriate indicators of a college's effectiveness in accomplishing its mission: (1) job placement rates of graduates, (2) retention and graduation rates, (3) ratings of the institution by graduates, (4) long-term outcomes of graduates and (5) quality of graduates' perfomance on the job (ACE, 1986). The ECS panel strongly encouraged the use of multiple methods of assessment to improve student and institutional performance, including standardized tests, student participation and completion rates, writing samples, senior projects, student satisfaction and placement, alumni and employer satisfaction, and faculty development and contribution (ECS, 1986).

Secretary Bennett has said that colleges should employ a variety of assessment methods, e.g., standardized tests, interviews, questionnaires, reviews of students' work, and studies of alumni and dropouts (USED, 1986). The National Governors' Association strongly recommends that institutions should use "a number of assessment approaches and techniques" because of the wide variety of intended outcomes set for undergraduate education and the diverse nature of the student population (NGA, 1986). Boyer warns that efforts to assess

student learning may "trivialize the B.A." unless colleges measure what is important. Most tests now focus on mechanics rather than measuring student abilities to communicate and think critically. He also states that assessments should deal with "the quality of campus life" and not just academic achievements. He recommends a required senior thesis, a senior seminar, a senior colloquium, a portfolio record of activities, and alumni surveys and interviews (Boyer, 1986).

Summarizing the recommendations of these reports with respect to the assessment techniques to be used, it would appear that achievement tests (standardized or locally developed), student persistence (retention) studies and alumni surveys would rank highest in terms of both popularity with the colleges and attracting the confidence of the publics external to the institutions. Much could be done to enhance national normative data from these three sources. In particular, cooperating groups of colleges and universities should be encouraged to collect and report these data (as group data) to interested parties, including a national statistics center.

Moving beyong the collection of descriptive data, however, there is a critical need for experimental data on what really works at the various institutions to improve institutional performance on such indicators as test score gains, student retention rates and alumni satisfaction indices. Cooperating groups of colleges and universities should be encouraged to collect base-line data on student performance (including alumni), then implement program changes, and then again measure student performance to determine if the experimental changes were effective. Such information would not only be helpful to the groups of institutions collecting the information, but also to other, similar types of institutions. In its survey, the American Council on Education found that at least 75 percent of the college presidents reported that a general review of curriculum within the past five years included general education requirements with a greater emphasis on such competencies as communicating and reasoning (ACE, 1986). One can only wonder how many of these institutions will have the data to know if these proposed changes are indeed improvements or just another round of changes.

Some of the reports cited above have suggested some program changes that could serve as a national agenda for experimental changes in programs and related data collection. For example, the NIE Study Group has recommended, (1) greater use of active teaching modes involving students in learning, (2) strengthen advising programs, (3) at least two full years of liberal education, (4) dissemination of a statement of the knowledge, capacities, and skills that students must develop prior to graduation, (5) greater emphasis on communication and reasoning skills in all liberal arts courses, and (6) supplementing the credit system with proficiency assessments as a condition of awarding degrees (NIE, 1984).

The Association of American Colleges has outlined a minimum required curriculum that emphasizes teaching students how to learn, suggests teaching communication skills in all core courses and urges colleges to create opportunities for more active learning in class, on campus and in the community (AAC, 1984). The Southern Regional Education Board has essentially endorsed the AAC report (SREB, 1985).

Boyer has recommended (1) an "integrated core" curriculum that emphasizes application of knowledge to life after graduation; (2) an "enriched major" to explore the history, social implications and value issues of the field; (3) increased exchanges between faculty and students in the classroom; (4) greatly expanded new student orientation and advising and (5) a clear set of goals for student learning at each institution (Boyer, 1986).

Summary

What is being suggested in this brief paper is a role for the Federal Government in the assessment of college outcomes that envisions groups of cooperating colleges and universities reporting results of their data collection efforts to the Federal Government as group data. Not recommended here is any additional reporting by individual institutions directly to the Federal Government. This will protect the confidentially of the individual institutions while providing data to the Nation through the Federal Government on the general state of American higher education as well as experimental efforts to improve undergraduate education.

It is anticipated that the cooperating groups of institutions will likely be formed by various types (large, small, public, private, two-year, four-year, etc.). They may correspond to existing State systems and/or private consortia. Funding for the cooperative data collection and reporting efforts may come from Federal, State or private sources. The reporting of such data to the Federal Government would be followed by dissemination of the data by the Federal Government to interested parties across the Nation.

The college outcome assessment data which the Federal Government should assist in collecting and disseminating should be that which the colleges and universities will find to be most critical to program improvement and that the general public will find to be most convincing in making appropriate judgments about the quality of American undergraduate education. In this light, three types of outcome data would appear to be most useful: (1) achievement test score gains, (2) student retention/persistence rates and (3) alumni satisfaction indices. The Federal Government should do what it can to assist and encourage the collection of these types of data by cooperating groups of colleges.

Lastly, it is suggested that the Federal Government actively promote a national agenda of undergraduate program changes that colleges and universities should implement on an experimental basis, with appropriate data collection procedures, to determine what strategies work best to increase institutional performance on the three indicators cited above.

References

AAC Association of American Colleges: A Search for Quality and Coherence in Baccalaureate Education, 1982.

ACC (Association of American Colleges); Integrity in the College Curriculum, 1984.

ACE (American Council on Education); Campus Trends, 1986.

Boyer, Ernest L.; College: The Undergraduate Experience in America, Carnegie Foundation for the Advancement of Teaching, 1986.

CFAT (Carnegie Foundation for the Advancement of Teaching); Missions of the College Curriculum, 1977.

COPA (Council on Postsecondary Accreditation); Evaluating Educational Quality, 1979.

ECS (Education Commission of the States); Transforming the State Role in Undergraduate Education, 1986.

Forrest, Aubrey; Increasing Student Competence and Persistence, The American College Testing Program, 1982.

Jaschik, Scott; "Public Universities Trying Tests and Surveys to Measure What Students Learn", The Chronicle of Higher Education, September 18, 1985, p. 1+.

Miller, Jerry W.; Credentialing Educational Accomplishment, American Council on Education, 1978.

NCEE (National Commission on Excellence in Education); A Nation at Risk: The Imperative for Educational Reform, 1983.

NGA (National Governors' Association); Time for Results: The 1991 Report on Education, 1986.

NIE (National Institute of Education), Study Group on Conditions of Excellence in American Higher Education; Involvement in Learning: Realizing the Potential of American Higher Education, 1984.

SREB (Southern Regional Education Board); Access to Quality Undergraduate Education, 1985.

USED (U.S. Department of Education); Assessment in American Higher Education, 1986.

The State of the States in Postsecondary Assessment

Edward A. Morante
New Jersey Department of Higher Education

Introduction

The issuance of "A Nation at Risk" in 1983 with its inflammatory language about rising mediocrity in our nation's schools aroused a rising tide of reform. It was little surprise that similar reports on higher education would follow. Reports from the National Institute for Education, The American Association for Higher Education, The Education Commission for the States, and the National Governors' Association are just four of the national reports which discuss concerns for undergraduate education and offer recommendations. Conferences, task forces, and blue ribbon panels, not to mention discussions in board rooms and faculty dining rooms have been held almost non-stop on this topic. Institutions of higher education have formed committees to study approaches to reform and to develop and implement assessment programs. Accrediting agencies have also begun to demand outcome measures in their assessment process, a development bound to attract the attention of those few remaining colleges who thus far have ignored the calls for change.

Many state agencies have also responded. In fact, according to a recent ECS (1986) report, over half of them have statewide initiatives to improve the quality of collegiate education.

After defining terms, this paper will provide an overview of a sampling of states leading the efforts in postsecondary assessment with special focus given to New Jersey. I will conclude with a series of questions that must be addressed if efforts to foster assessment among our colleges are to succeed.

What is "Postsecondary Assessment"?

In a recent AAHE bulletin (February 1986), Boyer and McGuinness discuss assessment:

> One of the difficulties with the term "assessment,"..... is that it increasingly is being used to refer to several different activities, from narrowly defined achievement testing to broadly defined student assessment, from evaluating student programs to evaluating institutional or program effectiveness, from diagnosing student needs to certifying minimum levels of student achievement.
>
> For the purposes of this paper, postsecondary assessment is not admissions testing nor is it raising admissions standards or admissions requirements. It is not working with secondary schools to improve the proficiencies of students entering collegiate education (no matter how important this is). And it is not curriculum reform or

incentive funding although these are frequently related. Rather, postsecondary assessment is collegiate evaluation that falls into three categories:

1. Developmental Education: This includes basic skills placement testing to assess students' proficiencies and readiness to handle college-level courses. In addition to assessing individual students, postsecondary assessment should include the evaluation of developmental education efforts. In this regard, focus can be either on process or outcome but should answer questions on the effectiveness of the developmental educational program.

2. College-level Learning: This includes "rising junior" examinations, sophomore assessment, and graduation tests. These are tests of cognitive proficiency to determine competency of students. They can be used to assess programs/institutions or can serve as "gates" to continued education, program admission, or certification licensure.

3. Outcomes Assessment: Assessment programs in this category are more holistic, examining more than student learning. Both cognitive (e.g. intellectual skills) and non-cognitive areas (e.g., affective and personal satisfaction and development) are assessed as well as post graduation activities (e.g., occupation, salary, graduate education). The impact and/or effectiveness of a college program, institution, or statewide higher education system are also possible areas of outcomes assessment.

Let me digress for a moment to differentiate outcomes assessment from traditional methods of evaluating colleges. In the past, most of the emphasis on evaluating colleges was found in student input variables (e.g., admissions test scores); faculty reputation (e.g., number and kinds of degrees, number of publications); size of endowment, amount of dollars in research and grants; number of books in the library; and physical facilities. The major change in postsecondary or collegiate assessment is to move away from these traditional indices to focus on the outcome or what impact a college has on students (e.g., the value added to what a student brings upon entry) or on some larger audience (e.g., a local community or a state).

State-by-State Perspective

Currently, statewide postsecondary assessment programs are at a relatively early stage of development, so that there is no such animal as a state model. Each state seems to be addressing (or not addressing) postsecondary assessment in its own way probably because of the unique collection of interacting variables and conditions prevalent in each state.

In addition, the definition of state role itself has multiple meanings. At one end of the spectrum, state governing agencies have imposed mandates on colleges to carry out specific assessment procedures. Other states are, or appear to be, requiring colleges to assess postsecondary education on several variables using specific instruments while leaving much of the assessment in the hands of local institutions. Still other states are merely encouraging assessment without much, if any, specification of what or how.

In this time frame, I cannot do justice to the variety of programs underway or being planned by all the states. Some states like Missouri and Wisconsin have within their borders colleges which have become national models of institutional assessment efforts (i.e., Northeast Missouri State University and Alverno College respectively). Other states, like Colorado or California, have recently enacted legislation which might lead to statewide initiatives and/or build upon programs already locally developed.

The following section offers a cross section of states which appear to represent the major efforts in postsecondary assessment in the United States today. New Jersey is treated subsequently and separately, not necessarily as the model for others to follow, but because it has been in the business of college assessment for a long time, has accumulated more statewide data, and, obviously, is the one with which I am most familiar.

Connecticut

The Connecticut Higher Education Coordinating Board has proposed a far-reaching system of statewide basic skills assessment using a single instrument at the college entry level. The Board also proposed a statewide college outcomes project using the same parameters across the public college system in the state. These proposals aroused much concern in Connecticut and the coordinating board backed off; it is now proposing that basic skills assessment and outcomes assessment be mandatory but allowing each institution to specify its own assessment methods and instruments. Each institution would be required to report annually to the coordination board on its locally developed assessment process. This program is still very much in the developmental stage with implementation currently planned for the 1989-90 academic year.

Florida

Created by the legislature in 1979 and administered for the first time in 1982, Florida has a rising junior exam that a student must pass to receive an associate degree or continue on to upper division courses at a four-year public institution. Although initiated by the legislature, the CLAST (College Level Academic Skills Test) was developed based on the identification by Florida college faculty of the essential skills that should be

possessed by all students by the end of their sophomore year. Reading, writing, and computation are the skill areas measured by the CLAST. Passing scores have been raised over the years and are scheduled to be raised again in the future. Students have three opportunities to pass the exam. Preliminary studies indicate that minority students have not performed especially well on the test. Projections have been made of a massive failure rate by minority students in the future unless present conditions change. Much controversy surrounds this testing program in Florida including strong opposition from Miami/Dade Community College.

More recently, Florida has addressed assessment of basic skills proficiency of entering freshmen. While no statewide exam exists in this area at this time, all public 2- and 4-year colleges are required to test their beginning students using one of four possible tests: ACT, SAT, ASSET, or MAPS. Students must meet statewide standards on one of these tests or enroll in remedial courses.

Georgia

There are two statewide testing programs for the higher education system of Georgia: The Basic Skills Exam and the Regents Test. Entering freshmen in the public colleges who score below 330 on either the verbal or mathematics sections of the SAT are required to take appropriate sections of the Basic Skills Test, a test developed in Georgia. While the SAT scores are statewide minimum cut-off scores, local institutions may set higher cutoff scores. Students who initially fail the skills exam must eventually pass it to fulfill their remedial requirements with or without remedial work.

The Regents Test, given to all students at the sophomore level, must be passed in order to receive a degree from either a 2- or a 4-year college. This test consists of two parts: reading and writing. Both sections must be passed to qualify for graduation. Students not passing either section of the Regents Test by the time they complete 75 credits must undergo appropriate remediation each quarter until they pass the exam. Transfer students are subject to similar requirements.

Rhode Island

Rhode Island has decided to develop its statewide assessment program by encouraging its three public institutions to establish self-assessment efforts at both the program and institutional level. A set of statewide quality indicators including admissions standards, retention and completion rates, placement information, and value-added data have been identified by the Office of Higher Education. Each college program in the State is required to report to the Board of Governors on a 3- to 5-year cycle the results of either an institutional level review or a national accreditation assessment. The emphasis is to use existing data and assessment mechanisms wherever feasible and to

organize such information into a comprehensive assessment package. Discussions have also begun on the possibility of instituting a common statewide basic skills placement exam for entering students. In Rhode Island, the role of the State appears to be that of a catalyst for local assessment within certain parameters.

Tennessee

Since 1979, Tennessee has had a performance-based funding program as an incentive for educational improvement at its public colleges and universities. The institutions are encouraged to define more clearly their objectives and to establish an assessment program to demonstrate how well these objectives are being met. The program was voluntary until 1984 but is now required of all public colleges. State guidelines are available to determine who qualifies for this additional, incentive budgeting.

As part of this effort, the ACT Comp test (an instrument designed to assess general education competencies) is required of all students who seek a degree. While no passing score is specified, the results are reported by institution and are used to influence each college's budget.

In addition, the State Board of Regents, which includes all public 2- and 4-year colleges except the University of Tennessee, began a statewide basic skills assessment program in 1985. In the first year, all entering students who scored below 16 on the ACT were required to take a statewide test in basic skills. The Descriptive Tests of Language Skills and Mathematics Skills (developed by ETS) were chosen by the Regents. Future plans call for testing all entering students regardless of initial ACT score.

The Tennessee Board of Regents has also begun a comprehensive statewide evaluation system of its colleges' remedial efforts. Focusing on outcomes indicators, the program will analyze data on such variables as passing rates, retention, credit ratio, pre- and post-testing, grade point averages, and graduation rates. A statewide computerized system tracks individual students as they progress and transfer through any college in the Regents system.

Texas

About a year ago, the Texas State Coordinating Board appointed a Committee on Testing (the "Hardesty Committee" after its chairman) to explore the development of statewide postsecondary assessment. After examining several possible models, the committee focused on two statewide systems, Florida and New Jersey, and sent task forces to these states for study and review. While the Coordinating Board has yet to act, the Committee's recommendations included the development of a comprehensive testing, placement, and remedial program at each public college. In particular, the committee recommended that

either the New Jersey College Basic Skills Placement Test be used or that a comparable test be developed for the State of Texas. Students identified as needing remediation would have to demonstrate competency before progressing to upper division courses.

New Jersey

New Jersey has two separate and distinct programs in post-secondary assessment: the Basic Skills Assessment Program (BSAP) and the College Outcomes Evaluation Program (COEP).

Basic Skills Assessment Program

The first program, established by Board of Higher Education mandate in 1977, has two main functions: to assess the basic skills proficiencies (i.e., reading, writing, and mathematics) of students entering colleges in the state; and to evaluate the character and effectiveness of the remedial/developmental programs at each of New Jersey's 30 public colleges. The Board also created a Basic Skills Council to carry out the program. Representing all sectors of higher education, the Council has served as a buffer between the colleges and the State Department of Higher Education.

Working with Educational Testing Service, the Basic Skills developed the New Jersey College Basic Skills Placement Test (NJCBSPT). There are two purposes of the test: to place students in appropriate beginning courses and to collect statewide data on the extent and level of entering student proficiencies. The NJCBSPT is a criterion-referenced power test composed of an essay and four multiple choice sections: reading comprehension, sentence sense, computation and elementary algebra. Board policy requires that all entering students must take the NJCBSPT, must be placed in remedial courses if needed, and must demonstrate proficiency before enrolling in college-level courses.

The NJCBSPT results are reported publicly each year and sent to high schools across the state. The results for entering freshmen have been rather disappointing:

26 percent fully proficient in verbal skills
28 percent fully proficient in computation
13 percent fully proficient in elementary algebra

There has been no significant change in these scores over the nine years of testing, although recent reforms in primary and secondary education in the state offer promise for improved results.

The second major function of the New Jersey basic skills program is the evaluation of the colleges' efforts to help their skills-deficient students. The Council recognized the complexity

of evaluating developmental education programs across 30 public colleges by a) focusing on outcomes (the colleges could provide their own brand of remediation) and, b) examining multiple indicators of effectiveness. Seven such indicators were ultimately decided upon:

1. passing rates in remedial courses;
2. retention rates;
3. pre- and post-testing;
4. performance in subsequent subject-related college courses;
5. grade point average;
6. credit ratio;
7. successful survival rate

The statewide results demonstrated a diversity of both programs and outcomes across the colleges. One finding was that at both the state colleges and the community colleges students who successfully completed remedial programs ("completers") are retained at a rate slightly higher than that of students who did not need remediation ("no need"). Specifically, at the nine state colleges retention rates for the completers was 75 percent as compared to 39 percent for those who did not complete needed remediation and 70 percent for the "no need" group. While grades among the completers were not as high as those not needing remediation, successful survival rates (a combination of retention and grades) were comparable. The results were similar for both a one-year and a two-year follow-up and have been replicated for several entering student cohorts.

College Outcomes Evaluation Program

In mid-1985, the New Jersey Board of Higher Education created the College Outcomes Evaluation Program and an Advisory Committee to plan and implement it. Included in the Board resolution was the development of a sophomore test to assess college learning beyond basic skills.

The COEP Advisory Committee divided its tasks into four subcommittees and charged each with a series of tasks:

1. Student Learning

The focus of this subcommittee is the assessment of how much and how well students learn in college. Specifically, the student learning subcommittee is charged with reporting on (a) the areas in which it is important to assess student learning; (b) the methods to assess this learning; and (c) the primary responsibility for such assessment (i.e., statewide, institution-based, etc). The primary emphasis of the subcommittee has been on institutional effectiveness, rather than the achievement of students as individuals. To date, the subcommittee has identified three areas for possible assessment: general intellectual skills (e.g., critical

thinking, oral and written communication and quantitative reasoning); a broad understanding in what has been traditionally labeled "general education," and knowledge and skills appropriate to a college major or concentration. The subcommittee has reached a consensus that assessment is needed and that it is appropriate to assess students at entry, at the end of the sophomore year (or an equivalent level), and before graduation. How many students should be assessed; which aspects, if any, should be standardized on a statewide basis and which should be designed locally; what specific areas should be assessed and what assessment techniques should be used are the issues of current focus. The subcommittee is also struggling with concerns about student motivation to participate in any assessment and the use of assessment results.

2. Student Development/Post-Collegiate Activities

This subcommittee has begun exploring the many outcomes of a college education that are not conveniently measured by traditional academic assessment procedures in the areas defined in the title. It has been debating the importance of numerous outcomes variables and plans to make recommendations about the feasibility and desirability of alternative assessment models.

Among the factors identified by the subcommittee for further review and definition are

- Student economic outcomes (e.g., student's first job earnings);
- Competence and skill measures (e.g., GPA);
- Morale, satisfaction and affective outcomes measures (e.g., alumni surveys);
- Status and certification outcomes measures (e.g., performance on certification exams);
- Retention (e.g., percent of students in a defined cohort who graduate in a particular academic program).

3. Research, Scholarship, and Creative Expression

Originally charged with focusing on research outcomes, the subcommittee has broadened its responsibilities (and title) to include both scholarship and creative expression, the related activities of faculty. This subcommittee has begun by developing a matrix of "first-order outputs" of faculty and staff activities. These include

- teaching;
- publications;
- inventions and technical developments;

- o performances and designs; and
- o professional advice, assistance, and participation.

The subcommittee is now developing a model linking these outputs to some of the desirable "outcomes" of research, scholarship and creative expression, and describing how they affect a variety of audiences such as students, peers, industry, the public, government, and the media. The members have reached consensus on the need to go beyond traditional approaches that focus on quantitative measures of faculty productivity such as numbers of publications and citations.

4. Community/Society Outcomes

The focus of this subcommittee is to determine how to describe and/or assess the impact a college or university as a whole has on the communities they serve, both narrowly and broadly defined. It is examining educational and cultural service as well as the economic impacts, both positive and negative, short- and long-term, that colleges and universities have on their local communities and on the state.

Using materials developed at NCHEMS (National Center for Higher Education Management Systems) as a starting point, the committee has identified a series of audiences and out-comes. The audiences include current students; former students; family, relatives, and associates of students; faculty, staff, and related individuals; interest-based communities (e.g., private enterprise, associations, government, public service groups, institutions); geographic-based communities (local, state, region, nation, the world); aggregates of people (e.g., populations defined by demographic characteristics; educational level, income level, occupation, sex, race, etc.).

The following broad categories of outcomes have also been identified:

- o Economic access and independence outcomes;
- o Economic costs and resources;
- o Economic costs and benefits;
- o Facilities and events;
- o Direct services;
- o Advisory and analytic assistance;
- o Treatment, care, and referral services.

In sum then New Jersey is attempting to create a comprehensive, holistic assessment program which will have both a statewide component and a requirement for local assessment at each college. Emphasis will be on using incentives including grants to encourage participation. Care is being taken to avoid a single model and to develop multiple indicators that take into account the diversity of missions and institutions across the state. The active and direct participation of over 60 individuals from the colleges, as well as the use of a network of

other college-based faculty and staff, and a planned series of statewide conferences emphasizing dialogue are expected to overcome much suspicion and concern about the state's assessment efforts.

Questions

In exploring the role of statewide efforts in developing post-collegiate assessment programs, it would be helpful to consider some of the questions that have been raised in New Jersey (and no doubt elsewhere):

- How can a statewide system be developed given the diversity of students missions, programs and institutions in the state?
- Will assessment improve higher education or harm it by diverting needed funding and energy or by focusing on narrow, overly simplistic, and misleading indicators of accountability?
- What should be the focus of assessment--individual students, programs, institutions, or the statewide system?
- What educational skills should students possess?
- Who should be tested or assessed? Who should do the assessing? When, and who should pay for it?
- How do you balance "value added" assessment with minimum competency?
- Should there be a single instrument or multiple measures? What should be assessed?
- Should a test be developed as a rising junior exam barring students from continuing unless they demonstrate appropriate competency? Are there other methods of assessing students' performance beyond paper and pencil tests? Are they feasible?
- What is the relationship between process and outcomes and which should be the focus?
- What data are already being collected and how should they be used?
- How do you balance access and quality?
- How do you balance accountability with goals of excellence?
- How can we balance a "top down" approach with a "bottom-up" or "grass-roots" approach?

- Can any system appropriately measure the essence of education, especially higher education, given the number and diversity of variables involved and the lack of control over these variables?

- Does college have a significant impact on students beyond normal maturation and outside influences and can this impact be determined?

Frankly, I'm not sure of the answer to any of these questions or whether they can be answered with certainty, but we've got to try. The complexities, the difficulties, the flaws should be recognized but should not prevent us from proceeding. Rather, the questions and concerns should encourage us to be flexible, cautious, and perhaps most important, holistic. Simple solutions will not work to describe collegiate education, may paint inaccurate pictures, and will probably be counterproductive to improvement and excellence. But neither can you wait until all the questions are fully answered to everyone's satisfaction. That's a formula for status quo and mediocrity. We must both work toward answering the questions as best we can, while continuing to develop comprehensive assessment programs.

While some institutions have and will continue to place a high priority on assessment, the complexity of the task, the competition of other pressing needs, and the discomfort of evaluation make it likely that states probably must play a role. Those of us in higher education must realize that the traditional answer of "Trust me" will no longer fly. While statewide efforts are sometimes scary, with justification, they can also be helpful, even necessary as incentives for constructive change. That's what we're trying to do in New Jersey--and I think we're going to succeed.

References

Association of American Colleges. Integrity in the College Curriculum. Washington, D.C., 1985.

Basic Skills Council. Effectiveness of Remedial Programs in New Jersey Public Colleges and Universities, Fall, 1982 - Spring, 1984. Trenton, Department of Higher Education, 1985.

Basic Skills Council. Results of the New Jersey College Basic Skills Placement Testing, Fall, 1984. Trenton, Department of Higher Education, 1986.

Boyer, C.M., and McGuinness, A.C. State Initiatives to Improve Undergraduate Education, Bulletin of the American Association of Higher Education, Feb., 1986.

Education Commission of the States. Transforming the State Role in Undergraduate Education: Time for a Different View, Denver, Colorado, 1986, National Governors Association, Task Force on Quality, John Ashcroft, Chairman, 1986.

National Institute of Education, Involvement in Learning: Realizing the Potential of American Higher Education, Washington, D.C., 1984.

Appendix B

List of Conference Participants

Adelman, Clifford
Office of Research, OERI

Alejandro, Frank Z.
Information Services, OERI

Ashwick, Helen
Center for Education Statistics, OERI

Baird, Leonard L.
University of Kentucky

Baker, Curtis
Center for Education Statistics, OERI

Bayless, David L.
WESTAT

Black, MacKnight
Center for Education Statistics, OERI

Bragg, Steve
Illinois Board of Higher Education

Brandt, Norman
Information Services, OERI

Brown, George H.
Center for Education Statistics, OERI

Carroll, C. Dennis
Center for Education Statistics, OERI

Carrallo, Sal
Office of Research, OERI

Cirino, Anna Marie
National Association of College and
University Business Officers

Davis, Lynda
Florida Department of Education

Drews, Theodore H.
Center for Education Statistics, OERI

Elfond, George
Educational Testing Service

El-Khawas, Elaine
American Council on Education

Engle, Penelope
Educational Testing Service

Faupel, Betsy
Center for Education Statistics, OERI

Forrest, Aubrey
American College Testing Program

Freund, William
Center for Education Statistics, OERI

Fry, Ray
Library Programs, OERI

Guillen, Robert L.
Association of Independent Colleges & Schools

Halstad, Kent
Office of Research, OERI

Jones, Paul D.
Council of Graduate Schools

Kaplan, Gail
National Conference of State Legislators

Kolstad, Andrew
Center for Education Statistics, OERI

Kopka, Tess
Center for Education Statistics, OERI

Lena, Marjorie
Council on Postsecondary Accreditation

Lewis, Gwendolyn
The College Board

Lord, Jerome
Office of Research, OERI

Ludwig, Meredith
American Association of State Colleges & Universities

Marchese, Ted
American Association for Higher Education

McDonald, Jean
National Governors' Association

McGeever, Jim
Program for the Improvement of Practice, OERI

Morante, Edward A.
New Jersey Department of Higher Education

Novak, Richard
American Association of State Colleges & Universities

Otte, Richard R.
Programs for the Improvement of Practice, OERI

Owen, Eugene H.
Center for Education Statistics, OERI

Peng, Samuel
Center for Education Statistics, OERI

Phillips, Gary
Center for Education Statistics, OERI

Rendino, Nilda
American Association of State Colleges & Universities

Schmidtlein, Frank
National Center for Postsecondary Governance & Finance

Showalter, Stuart W.
Council for Advancement of Private Higher Education

Song, Tongsoo
Center for Education Statistics, OERI

Stacey, Nevzer
Office of Research, OERI

Stern, Joyce
Center for Education Statistics, OERI

Sumberg, Alfred D.
Florida Department of Education

Sweet, David A.
Center for Education Statistics, OERI

To, Duc Le
Office of Research, OERI

Treacy, Maureen E.
Center for Education Statistics, OERI

West, Joel D.
Office of Postsecondary Education, U.S. Department of Education

White, Lyn M.
National Endowment for the Humanities

Williams, Mary
Center for Education Statistics, OERI

Wittstruck, John
State Higher Education Executive Offices

Woods, Jacqueline
American College Testing Program

Appendix C

Biographical Sketches of Speakers

Leonard L. Baird has over 20 years experience assessing students and institutions in higher education. After receiving a doctorate in educational psychology at UCLA, he worked as a researcher at the American College Testing Program and at the Educational Testing Service, conducting studies in a variety of aspects of higher education. He is a professor of higher education at the University of Kentucky.

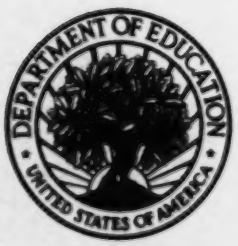
Paul R. Pintrich holds a dual appointment at the University of Michigan, Ann Arbor. He is an assistant research scientist at the University's National Center for Research to Improve Postsecondary Teaching and Learning. He also holds a half-time appointment as a lecturer in the University's School of Education. His doctorate in education and psychology is from the University of Michigan.

Aubrey Forrest is Director of Instructional Design and Assessment at the American College Testing Program. His principal responsibility is the management of the College Outcome Measures Program (COMP). He has directed more than 60 State, regional, and national seminars on outcome assessment and is the author of more than 50 articles, books, and test instruments.

Edward A. Morante is Director of the College Outcomes Evaluation Program, New Jersey Department of Higher Education. He was formerly Director of the New Jersey Basic Skills Assessment Program which was a state-wide effort to evaluate remedial-developmental education programs. He holds a Ph.D. in counselling psychology from Teachers College, Columbia University.

Dennis P. Jones is president of the National Center for Higher Education Management Systems and has over 25 years of experience in higher education. His areas of expertise include the assessment of the condition of institutions and state systems, and assessment of the needs of industry in a region or State. He has served in an advisory capacity to the Federal Government many times. Mr. Jones holds the M.S. degree in management engineering from Rensselaer Polytechnic Institute.

•U.S. GOVERNMENT PRINTING OFFICE:1987-193-046



CS 87-445